MODERN RADIO TECHNIQUE

General Editor: J. A. RATCLIFFE

PRINCIPLES OF RADAR

PLATE I

A plan position indicator display. The grid squares are of 10 km. side, and the circles of radii 10, 20, 30 and 40 miles around the station, Heath Row. The mass of echoes within the 10-mile circle, and those south-eastward at about 14 miles range, are almost certainly ground clutter, and the radial lines in the east-south-east direction are time-bases brightened by interference. The remaining detached arcs of brightness represent aircraft echoes.

PRINCIPLES OF RADAR

BY

DENIS TAYLOR, PH.D.,

Superintendent, Telecommunications Research Establishment,
Ministry of Supply

AND

C. H. WESTCOTT, PH.D.,

Physics Department, Birmingham University;
Sometime Senior Scientific Officer, Telecommunications
Research Establishment, Ministry of Supply

CAMBRIDGE
AT THE UNIVERSITY PRESS
1948

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
(Cambridge, and Bentley House, London)
Agents for U.S.A., Canada, and India: Macmillan

PREFACE

The present volume is addressed to those engineers, physicists and mathematicians who have some knowledge of the fundamental principles and pre-war practice of radio and who wish to learn about radar. It should also be of service to advanced students of radio, as well as to many who have worked on some aspect of radar but who wish to see the subject presented as a whole. A knowledge of electromagnetic theory, the calculus, and the $j\ (=\sqrt{-1})$ convention is assumed, as well as some acquaintance with normal radio practice. The present book, as its title indicates, deals with the principles of radar. Details of radar techniques are to be dealt with in other volumes of this 'Modern Radio Technique' series, which will include books on wave guides, circuit technique, etc. We have also excluded from this volume the whole field of radio navigation, which is the subject of another volume in this series.

The first eight chapters of the present book are therefore devoted to an exposition of the principles of radar, excluding those cases where a 'responder' equipment is placed in the target to be located. The uses and principles of responder methods are dealt with briefly in Chapter 10. Chapter 9 gives a summary of the characteristics of a selection of typical practical radar equipments; it is thought that this will be of value in correlating the references in earlier chapters to particular applications of the principles, as well as illustrating how factors such as the choice of wave-length work out in practice. The reader's attention is also directed to Appendix 1, which contains a collection of standard formulae together with a list of units and symbols. It will be seen, for example, that throughout the book E always represents a field strength, V being always used for a voltage and v for a velocity. We have also used the term 'sender' throughout rather than 'transmitter', regarding 'transmission' as the act of propagating a wave and not of originating one.

We are aware that in a few matters there is room for an honest difference of opinion. For example, others may not follow exactly the definition of 'radar' that we have adopted in Chapter 1, but such differences are matters of convenience rather than of principle. In one or two other places, notably concerning aerial noise, we touch

on subjects where research is still active and may completely change our ideas. We have, for example, made no mention of solar noise, which, at the present stage of the 11-year cycle, is becoming increasingly important; on occasions of solar activity aerial noise may considerably exceed the figures we have mentioned.

The authors wish to acknowledge the assistance received in discussion and advice from friends and colleagues in the Government Service and elsewhere. They are especially indebted to Mr L. H. Weaver who assisted in the preparation of the line drawings. All the half-tone plates are Crown Copyright, and the courtesy of H.M. Stationery Office in allowing their reproduction is hereby acknowledged. The authors are also indebted to the Ministry of Supply for permission to publish this book.

DENIS TAYLOR
C. H. WESTCOTT

May 1947

CONTENTS

APPENDICES

LIST OF PLATES

Chapter 1

INTRODUCTION

Radar, or radiolocation (the terms are interchangeable), will appear to most people to be a development of the war which has just ended. In fact, however, the methods used date from as long ago as 1924, when the height of the ionosphere was first measured. As a military weapon the developments began in 1935, and at the outbreak of the war in 1939 a few radar stations were already operating. Secrecy restrictions naturally prevented this fact being generally known, and detailed technical information about radar has only been released to the general public since the end of the war. It is therefore not surprising that the whole subject should be thought of as a wartime development, even if its potentialities for peacetime are also recognized.

This book is intended as an exposition and survey of the principles underlying radar design, and it is hoped that it will be of service to readers who have found the sudden release of so much detailed information to lead to mental indigestion, by indicating the common factors which underlie the many types of radar equipment which have been developed and used during the war. In a book of this size, detailed consideration cannot be given to every item of the equipment; other monographs of the present series will attempt to deal in greater detail with specialized aspects of the subject such as micro-wave technique and receiver and presentation system circuits. This volume will be largely concerned with the factors determining the performance of any particular radar equipment, and those determining its suitability for a specified purpose, rather than the detailed behaviour of its component parts.

It will be convenient to define radar or radiolocation before proceeding further. Both Appleton* and Watson-Watt† have given definitions, and following them we may say that, briefly, radar is a system for measuring at least two of the coordinates which define the position of an object with respect to the observer: one of these

* *J. Instn Elect. Engrs*, vol. 92, p. 340 (1945).
† Ibid. vol. 93, p. 374 (1946).

coordinates is the distance between them and is inferred directly from the time of travel of a radio wave travelling by the shortest path between them. Watson-Watt distinguishes between primary radar, where the object to be located plays the purely passive part of reflecting a radio wave, secondary radar in which a responder is used to provide the return signal, and radio navigation in which, in principle, no return signal is necessary. He also extends his definition to cases where the distance determined from a time delay is the difference between the paths of two radio waves, but for primary radar this extension is unnecessary. In this book we shall generally restrict our attention to primary radar as defined by Watson-Watt, although many of the principles are common to the two systems, and an understanding of the more complex two-way transmission problems of primary radar should lead without difficulty to an understanding of the simpler cases of one-way transmission to or from a responder.

From this it should be clear that classical direction-finding is not radar within our definition, although a position may be inferred by bearing cuts from two different observing stations. This would be true even if the measurements were made, not on transmissions from the object being located, but on radio waves scattered by this object. Classical direction-finding methods may be and, with suitable modifications on account of the different wave-length employed, frequently are used as part of a radar equipment to give information additional to the range, but it is the measurement of range directly that is the essence of radar. The velocity of light (and all other electromagnetic radiations) is, of course, considered as a known fundamental constant of this measurement.

The first measurement of the distance travelled by radio waves was made in 1924 by Appleton and Barnett,* who used a frequency-modulation method. While not within our definition of radar, since no second coordinate was measured, this was essentially a radar method. The radio-frequency of the sender was increased linearly in a time τ, from f_a to f_b (fig. 1·1), but the wave returned from the ionosphere (at a distance of r, say) is delayed by a time interval $t = 2r/c$, as shown by the dotted line in the figure. The latter wave therefore differs in frequency from the wave being transmitted at the same time, and 'beats' result. If $B = f_b - f_a$ is the band swept

* *Proc. Roy. Soc.* A, vol. 109, p. 621 (1925).

in time τ, the difference between the two radio-frequencies, and therefore the beat frequency, is given by

$$\frac{t}{\tau} \times B = \frac{2Br}{\tau c},$$

which is seen to be proportional to the range r of the ionosphere. The total number of beats resulting from one sweep across the frequency band is this frequency multiplied by τ, i.e. is $2Br/c$,* so that the range can be determined simply by counting the number

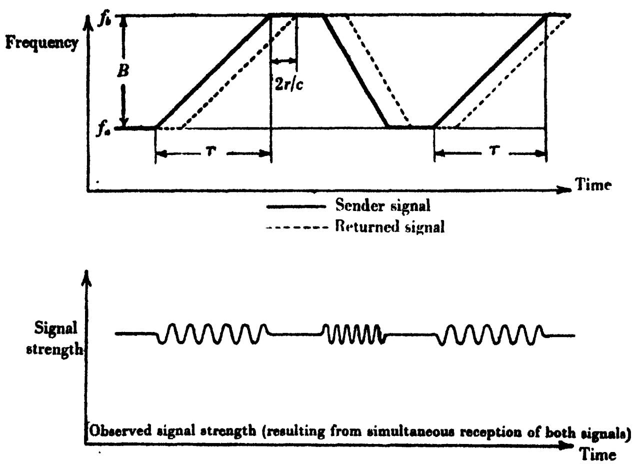


Fig. 1·1. Illustrating the frequency-modulation method.

of beat cycles per sweep, providing only that the sender frequency remains constant at each end of the sweep for at least a time $2r/c$ as shown in fig. 1·1. This figure also shows a typical record of the received intensity on which the beats are exhibited.

In 1925 Breit and Tuve† developed an alternative method using amplitude modulation. In this method a pulse of radiation short

* This result is also obtained by taking the difference between the number of wave-lengths in the path travelled ($2r$) at the extreme frequencies, i.e. $2r/\lambda_b - 2r/\lambda_a$, remembering that

$$B = f_b - f_a = c/\lambda_b - c/\lambda_a.$$

For every wave-length change in the delay path one beat cycle results.

† *Phys. Rev.* vol. 28, p. 554 (1938).

compared with $2r/c$ was transmitted, and the reflected pulse observed directly at a time $2r/c$ later. It was usual to display the reflected pulses on a cathode-ray tube time base, and deduce the delay, and hence the range, from the distance the spot had moved across the screen. Fig. 1·2 illustrates this method. The particular advantage of the pulse method, which has led to its very general adoption in preference but not to the exclusion of the frequency-modulation method, arises only when more than one reflecting object is present. Whereas the pulse method gives two or more clearly resolved echoes from objects at different ranges at the corresponding points on the time base, the frequency-modulation method gives only a complicated 'beats' wave form requiring Fourier analysis to separate the components.

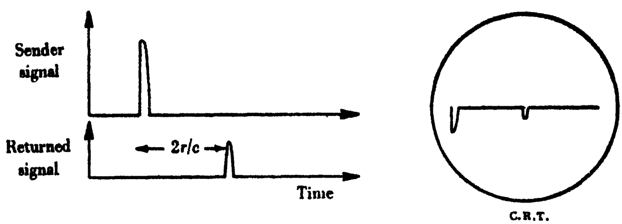


Fig. 1·2. Method due to Breit and Tuve.

These methods were developed considerably before it was realized that objects smaller than the ionosphere could give recognizable echoes. In 1931 certain Post Office engineers* observed the reflexion of radio waves from aircraft in the vicinity, but they regarded this as 'interference' and failed to realize its possibilities. In 1933 similar results, with a more complete survey of the possibilities arising from such reflexion, were published in America,† and in 1935 work was begun on radiolocation proper at an Air Ministry station, the whole matter promptly disappearing under a cloak of secrecy. It was realized that the problem had become a technological one of obtaining higher power senders and more sensitive receivers in order to obtain adequate range, no new principles being involved. It is the tremendous advances in such matters of technology which have recently been published for all to read. The

* *Post Office Rep.* no. 223 (June 1932).

† England, Crawford and Mumford, *Proc. Inst. Radio Engrs, N.Y.*, vol. 21, p. 464 (1933).

magnitude of the problem has been expressed by Watson-Watt, who pointed out that the radar location of an aircraft at reasonable range was early seen to involve sender powers of the 100 kW. order, with receivers sensitive to a micro-micro-watt, so that in the engineering sense radar has an 'efficiency' of the order of 10^{-17} only. All the rest of the transmitted power is dissipated uselessly or lost in space.

The importance of radar during the war has been amply recognized in newspaper and magazine articles and has fully justified the large sums expended. Here we can only mention a few highlights—enabling our limited fighter resources to be in the right place at the right time during the Battle of Britain; night fighter aids in the Night Battle which followed and in the North African campaign; radar for detecting surfaced submarines* in the anti-U-boat war; radar bomb-sights in the bombing of Germany; anti-aircraft and other aids against the V 1 flying bomb; the sinking of the *Scharnhorst* and many other naval engagements. It is also clear that radar will have many uses in peace for both aircraft and shipping. By its means ships will be able to steam at full speed through the thickest fogs, to determine their positions accurately when near coasts and to steer through narrow channels, as well as obtaining warning of the position of other craft and icebergs. Aircraft can also be fitted with collision-warning devices, and in dense traffic areas, such as near airports, can be controlled from the ground from radar information. Blind-landing devices will also be of great value, and will largely use radar methods, while pulse navigational aids will be of use to aircraft and shipping alike. On land radar may also find application, although it is more difficult to predict probable developments here. But in any case it is clear that radar techniques will play a considerable part in our future civilization.

We may here conveniently resume our brief survey of the historical development of radar, starting from 1936 and dealing with matter only recently released from secrecy. The early radar equipments in England were designed to detect aircraft approaching our coasts, and used wave-lengths in the 6–15 m. band. They required large and expensive towers, the transmitting towers being generally 360 ft. high and the receiving towers 240 ft. These towers were

* It should be appreciated that radar is only able to locate a submarine if some part of it is above the water-line.

necessary to obtain a satisfactory radiation intensity at elevations below 10°, aerials some 10 wave-lengths above the ground being essential for this purpose. A peak power of 200 kW. was used, a pulse being transmitted every $\frac{1}{25}$ sec., and ranges of the order of 140 miles (depending to some extent on the characteristics of the site) were obtained on aircraft flying at 15,000 ft. A workable system on these lines was available at the outbreak of the war, and the chain was rapidly extended round the south and east coasts, and less rapidly round the west and north. Minor improvements were made later, and some stations were equipped with 1 MW. senders but fundamentally this system remained unchanged from the type existing in 1939. The main line of development lay at shorter wave-lengths: the gun-laying equipments in the 3 m. band, the coastal chain for the detection of low-flying aircraft and the ground control of interception stations on 1·5 m., early airborne 1·5 m. equipments, and the very great developments in all fields on wave-lengths of 10 cm. and below which became general in 1942. Naval developments showed the same tendency, and latterly the centimetre-wave equipments became predominant for all functions. Magnetrons were developed to give powers of the 200 kW. order at 10 cm. for airborne applications, and 1 MW. 10 cm. magnetrons were later produced for ground stations where a greater bulk and weight of equipment were allowable. Ranges of 8–10 miles on aircraft from aircraft thus became possible, as well as high-resolution ground stations of reasonable range, although the longer wave-lengths still retained their usefulness in ground equipments—1·5 m. coastal chain stations, for example, frequently plotting aircraft at 150 and sometimes at 200 miles and over.

The reasons for this downward trend of wave-length will become clear as this book proceeds, but it should be emphasized at the outset that while the drop to 1·5 m. and even 50 cm. was a normal type of development following more or less conventional lines, using, for example, triode oscillators in the sender and pentode r.f. amplifiers in the receiver, the work at 10 cm. and below was essentially a new fundamental development starting afresh from the properties of the electron. This pioneer work, undertaken as a piece of almost pure scientific research, has paid dividends of the highest order and led to a veritable revolution in radar. Beside it other developments, such as the plan position indicator (in which the radar draws its

own map on a cathode-ray tube showing directly the positions of all the objects detected) fade into relative insignificance.

While, therefore, radar was extending from the thin red line of a coastal chain inland to include gun-laying and intercepting fighter control stations, seawards to ship-borne radar, and skywards to the many airborne devices, the new centimetric methods were being developed. For land and sea radar they gave increased accuracy and improved the plan position indicator almost out of recognition, while for airborne radar, where small size of aerial was of prime importance, they led to the first really adequate answers to all the problems of air operations—the A.I. ('Air Interception') enabling fighters to locate the enemy bomber in the night sky, the A.S.V. ('Air to Surface Vessel') for anti-U-boat patrol, and finally the H_2S which provided the bomber pilot over Germany with a map of the country over which he was flying, in darkness and through cloud, so that he could drop his bombs where they were intended to fall.

We may note here that, in general, the radar station may be land-, ship- or airborne, and the target reflecting the radiation may be a land feature (e.g. a cliff or a church spire), a ship, an aircraft, or a surfaced submarine; echoes are also received from the waves of the sea, and one of the major problems of radar is the detection of a wanted echo among unwanted echoes such as sea or land clutter.

The essence of radar is the timing of an echo delay $2r/c$. Here c is the velocity of electromagnetic radiation, known to be $2{\cdot}9977 \times 10^8$ m./sec., in free space, but when high accuracy is required some correction is needed for the refractive index at radio-frequency of the atmosphere. It varies with meteorological conditions in a manner which will be considered in detail later. The correction, however, is generally small, being of the order of 1 in 10,000, but in some cases bending of the rays due to refractive index gradients becomes quite important.

The method of timing the delay may, as has already been indicated, employ either frequency modulation or pulse transmission. The former is less often used, being normally restricted to single-ranging devices, but is very useful in some applications, e.g. a radio-altimeter for measuring the exact clearance of an aircraft from the ground beneath. Fig. 1·3 shows a schematic of such a frequency-modulation system. The modulation may be mechanical (e.g. a

rotating condenser) or electronic, and is synchronized with a recorder which may take the form of a beat-cycle counter which resets after each sweep and passes a smoothed reading on to a meter display. It is essential that some of the sender power shall reach the receiver direct, but the signal so received must not 'swamp' the receiver. In the radio-altimeter it may be convenient to use sender and receiver aerials which are fairly directive and beamed downwards, one located on either wing, so that the two received signals are not too disproportionate in size. The accuracy of such a frequency-modulation system generally corresponds to ± 1 beat cycle, i.e. an error in range of $\pm c/2B$, so that if an accuracy of 1 m. is required, B must be 150 Mc./sec. Such a wide frequency sweep requires the use of a centimetre radio-frequency, say around 3000 Mc./sec., and is then quite feasible.

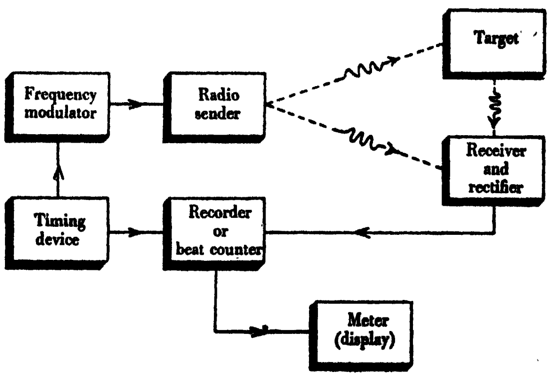


Fig. 1·3. Schematic of frequency-modulation system.

The pulse method, using amplitude modulation, is by far the most common radar system. Fig. 1·4 shows a schematic for this system. A multivibrator, master oscillator, or rotary device fixes the pulse recurrence frequency and feeds the modulator and presentation time-base systems. The modulator causes the radio-sender to transmit a pulse of preset duration and form, which is sent via the aerial and reflected by the target to the receiver aerial. In some developments the same aerial is used for both sending and receiving, using an automatic switching system. The receiver has to have suitable pass band characteristics to deal with the pulse modulation side bands, and must be free from paralysis following reception of

a large signal (e.g. that direct from the sender). Its output is rectified and amplified and applied to the cathode-ray tube display, either as a deflexion or an intensity modulation. The cathode-ray rube is also supplied with a range time base as shown, synchronized to the pulse transmission, and suitable calibration devices (not shown) are usually also necessary. The form of the display unit depends on the particular radar application, and on what coordinate or coordinates other than range are to be presented. For example, in the plan position indicator the time base starts from the centre of the tube and goes radially outwards in a direction depending on the azimuth of the beamed aerials (which sweep continuously), and the signal is presented as a spot intensifier. Many other forms of presentation are also in use for various purposes.

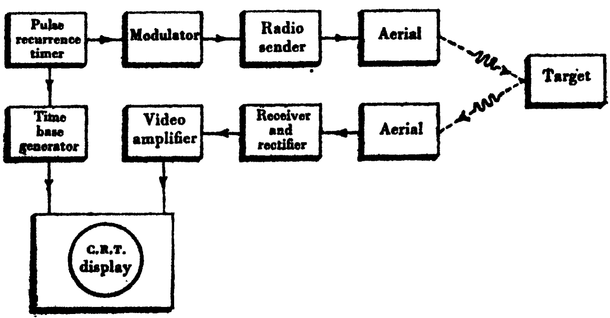


Fig. 1·4. Schematic for pulse radar.

Further detailed consideration of the pulse method will be given in the following chapters, but it should be pointed out here that there is a definite minimum range for the pulse method. Neglecting any additional effects due to receiver paralysis, the range cannot be measured unless the echo pulse does not begin until the sender pulse is finished, so that if the pulse length is τ, the minimum range is never less than $\frac{1}{2}c\tau$; this corresponds to 150 m. for a 1 μsec. pulse. In contrast, the only theoretical limit to the minimum range for the frequency-modulation method is that the range shall be greater than the error $(c/2B)$, although in practice the limit is usually a little larger than this. For the altimeter considered above, for example, the minimum range is theoretically only 1 m. ground-clearance, and in practice might be 2–3 m.

Chapter 2

THE GENERATION AND RECEPTION OF PULSE-MODULATED SIGNALS

In this chapter we shall consider in more detail the problems of the generation and reception of pulse-modulated radar signals. For the purposes of this chapter we shall have to suppose the radio-frequency f_0 to have been already selected, as well as the pulse-recurrence frequency f_p, and the pulse length τ. In anticipation of what will be explained later, we may say that the maximum pulse-recurrence frequency may be determined by the mean dissipation limits of the radio-generator, but that more often it is determined by the maximum range from which echoes are required, the spacing of successive pulses $1/f_p$ having to be at least some numerical factor greater than the longest echo-delay $2r_{\text{max.}}/c$. The pulse length may also be limited by mean dissipation considerations, but is otherwise set from the fact that the resolution of the system in range is $\frac{1}{2}c\tau$ (cf. end of Chapter 1), and the range accuracy attainable some numerical factor smaller than this, usually about $0{\cdot}1c\tau$.

It will also be clear from the previous chapter that the requirement for the generator valve in the sender will generally be as high a peak-power output as possible. The requirements* for the receiver are the very maximum sensitivity coupled with the ability to pass adequately the pulsed signals. The latter involves necessarily a rather wide pass band. The sensitivity is not limited in general by the available amplification, but by the noise level at the input stages of the receiver, and the fact that a wide pass band is necessary increases the noise level. Thus we have to push everything as near to the ultimate physical limits as possible, and the considerations of design have to be taken further than is general for communications equipment. It should also be noted that at the frequencies used in radar the aerial noise is generally small compared with receiver noise, so that it is absolutely essential that the latter be reduced to the very minimum obtainable—a consideration which does not arise at longer wave-lengths.

* The receiver must also be free from paralysis following the reception of a large signal (cf. Chapter 1).

2·1. Pulsed generators up to 1000 Mc./sec.

Up to radio-frequencies of about 1000 Mc./sec. it is usual to employ oscillators of conventional type, generally triodes, for generating the radio-frequency signal. Owing to the inertia of the electron it becomes increasingly difficult to generate appreciable power as the radio-frequency increases. The design of oscillator valves for pulse working differs somewhat from that for c.w. operation, and in particular the fact that the mean dissipation is lower for a given anode voltage may enable valves of physically smaller dimensions to be used, thus lowering the transit time and enabling us to push the frequency limit up a little. On the other hand, a high cathode peak emission is required, so that we may find a relatively fat cathode in a valve whose anode cooling arrangements appear quite insufficient by comparison with normal c.w. practice.

The power generator valves in the early 12 m. radar were totally enclosed in a silica envelope, with an 800 W. pure tungsten filament and a 1 kW. c.w. anode rating, and gave pulse powers of 20–30 kW. Later developments included the use of air-cooled copper anodes and thoriated filaments; these valves would give powers of about 100 kW. per pair at 200 Mc./sec., while at 12 m. in a power amplifier arrangement, a total pulse power of 1 MW. was attained. At 1·5 m. wave-length the use of a pair of self-oscillating triodes as the generator was universal, but at the longer wave-lengths tetrodes were also used, and an oscillator followed by a separate power-amplifier stage became a common arrangement. All these valves were of fairly large size, similar to conventional c.w. tubes, and 1·5 m. was roughly the frequency limit with this construction; even at this frequency the tuned circuits associated with the valves had become reduced to quite short lengths of transmission line, although at 12 m. it remained normal, with separate inductances and tuning condensers. Many different types were developed for different purposes, for a summary of which the paper by Ratsey*, and references given therein, should be consulted.

The most radical development in this field of 'conventional' triode oscillators was the development of the 'micropup' series of valves, which were particularly suitable for use with transmission-line circuit elements at relatively short wave-lengths. The first of these had an anode diameter of only about 1 in., with fins for forced

* *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 246 (1946).

air cooling and a thoriated filament consuming some 60 W. and would operate down to a wave-length of 1 m. Later a revised design incorporating an oxide-coated cathode, consuming only about 36 W., was produced, and a pair of these valves could be made to give 200 kW. peak at 50 cm. wave-length. These valves were later adopted for some 200 Mc./sec. equipments, in addition to their use at higher frequencies, and enabled a much lighter and less bulky transmitter to be built for this frequency than was possible with the earlier valve types.

For this series of valves the electrodes were arranged coaxially and the connexions to the circuit elements made in such a form as to preserve as far as possible the transmission-line characteristics. In fact the upper frequency limit of operation was often determined by shrinkage of one of the external circuits to negligible dimensions, only that part within the glass envelope of the valve being used. For this reason the actual valve circuits would often be unrecognizable by a person familiar only with pre-war technique, i.e. lumped L and C elements or relatively well-defined line elements.

2·1·1. *Centimetre band power generators*

When we come to the centimetre wave-band (3000 Mc./sec. and above) completely new techniques are necessary. This is a specialist subject and we can only outline the main developments here. First came the low-power oscillators such as are used as receiver local oscillators. These are of various types (Heil tube, double-cavity klystron, and reflexion klystron), but all make use of the principle of velocity modulation of the electrons. Instead of using a grid to modulate the density of the electrons (as in the conventional triode) and finding electron inertia a limit on the frequency at which we can operate, we use the inertia of the electron—we use one electrode to modulate the electron velocities, so that some go faster and others slower; then we let them travel for some time so that the faster electrons catch up the slower and the electrons are bunched. If we then collect these electrons on an anode we shall obtain bursts of current which we may arrange to be so phased that they maintain an oscillation in some resonator (usually a cavity) which forms part of the tube.

For a power generator the cavity magnetron is in universal use. This is shown diagrammatically in fig. 2·1. It consists of an anode

split into a number (6–12) of sections, between each of which is a built-in resonant cavity. There is a concentric cylindrical cathode, and the whole is placed in a magnetic field parallel to the axis of the cylinders, so that the normal path of an electron is not a radial line from cathode to anode but a slow spiral outwards. The action of such a magnetron is quite complicated, but in general we may say that if the cavities are excited, the electrons as they pass the mouths of the cavities are velocity modulated, and subsequently 'bunch' as they travel along. The curvature of their tracks in the magnetic

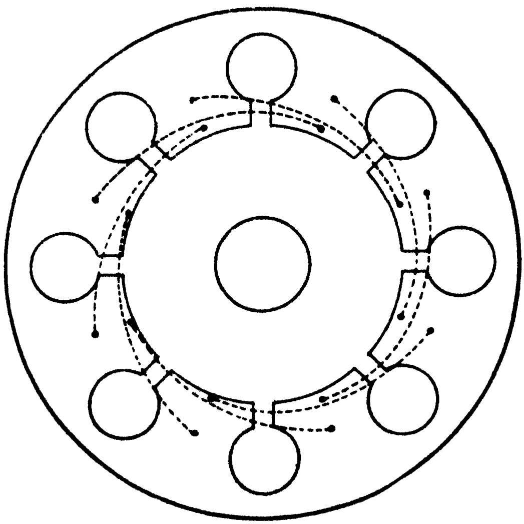


Fig. 2·1. Section of a cavity magnetron. (Dotted lines illustrate the method of strapping.)

field depends on their energy, and it is so arranged that they strike the anode in bunches at such a position and phase as to build up the oscillations of the cavities. The operation is much complicated by space-charge considerations; many of the electrons never reach the anode but are eventually bent by the magnetic field back to the cathode, and those that do reach the anode spend much longer on the way than they would in a normal diode, so that the space-charge density is very high by normal standards. Several modes of oscillation of the cavities are possible, but the one generally used is that in which alternate cavities are in anti-phase, so that alternate segments of the anode have positive and negative r.f. voltages on them

at any one instant. The device of strapping together alternate segments, as shown (dotted) in fig. 2·1, is most useful in suppressing unwanted modes and improving over-all efficiency. Power may be abstracted from the system by a probe or loop coupled suitably to one of the cavities.

At first (mid-1940) such magnetrons only gave an output of 5–10 kW., but this was later stepped up to 50–100 kW., and for ground equipments, where weight and bulk were not restricted as they were for airborne equipments, to 1 MW. These improvements were purely improvements in design, no fundamental innovations being involved. The frequency was pushed up in parallel developments, and the 10,000 Mc./sec. band is now quite as much a matter of routine as the 3000 Mc./sec. band. Very considerable developments have occurred at even higher frequencies. It will, of course, be realized that a given magnetron has its operating frequency fixed within narrow limits when it is constructed, since the resonant cavities form part of the vacuum system. The frequency may be variable over perhaps $\frac{1}{2}$ % with change of operating conditions, and specially constructed tunable magnetrons may be able to be varied over as much as a 5 % frequency range, but in general one has to buy a magnetron for use on a given frequency only. It may also be remarked that the use of relatively high 'Q' cavities in the magnetron limits the speed of modulation, and it is only because the radio-frequency is so high that this is allowable. Thus for a 3000 Mc./sec. cavity with a Q of 1000, we are limited to a modulation time constant of 0·1 μsec., which we are unlikely to need to exceed, whereas the same Q at 250 Mc./sec., say, would give a time constant of 1·3 μsec., which would be a severe limitation.

2·2. The modulator

The modulator system is equally important for pulse radar with the r.f. generator—relatively more important, in fact, than for orthodox modulated c.w. working, since the modulator has to handle large peak currents, and it is therefore frequently bulkier than the low mean-power oscillator with which it is associated. The rapidity of the modulation required also tends to add to the size of the modulator. The early long-wave systems used grid modulation, the oscillator grids being held some 1 kV. negative between pulses, but valves of considerable size were needed as modulators. In the

4 m. gun-laying and 1·5 m. equipments it was arranged that the modulator valve switched this bias off, but the end of the pulse was determined by a grid-squegg action, an anode-squegg being added to assist this process. With the advent of oscillator valves with oxide-coated cathodes it became necessary to employ anode or series modulation to avoid having a high d.c. potential applied continuously to the anodes. This was desirable to avoid 'flash-arcing' within the valves, which was particularly likely in the small-clearance valves produced for high-frequency operation. Applying the voltage for the duration of the pulse only, enabled much higher anode voltages to be used with safety. As series modulators, thyratrons were first used, spark gaps being later substituted; the latter could either be made to break down by superimposing a 'trigger' voltage pulse, or rotary gaps could be used, in which the spark struck when the rotating electrode approached the fixed electrode. In all these cases an 'artificial line' pulse-forming network was used in addition to the arrangement for initiating the pulse, voltage doubling types of circuit being favourites.

Modulators for centimetre magnetrons were also series devices, the anode being kept at earth potential and a negative square-wave voltage being applied to the cathode by the modulator. Since the action of the magnetron depends somewhat critically on the anode voltage it is important that the wave shall be square, i.e. that the voltage applied shall be constant throughout the pulse. Several methods of modulation have been used, including the thyratron and the triggered enclosed spark gap, in each case in conjunction with a pulse-forming network. For further details the paper by Ratsey cited in § 2·1 above, and references contained therein, should be consulted.

2·3. The Fourier components of a pulse-modulated signal

Before we can proceed to consider the behaviour of a receiver to pulsed signals it is necessary to examine the spectrum of such signals. The modulation wave form is periodic with a fundamental frequency f_p (fig. 2·2), so that its spectrum must consist of a number of harmonics of this frequency. Since the pulse length τ is much smaller than the pulse interval $1/f_p$, much of the energy will lie in the region of very high harmonics, having frequencies of the order of $1/\tau$. For example, f_p may be 500 sec.$^{-1}$ and $\tau = 2\,\mu$sec., in which case much

of the energy will lie in the region of the 1000th harmonic. If the modulation function is composed of square pulses, as in fig. 2·2*b*, the Fourier component intensities are readily calculable; the

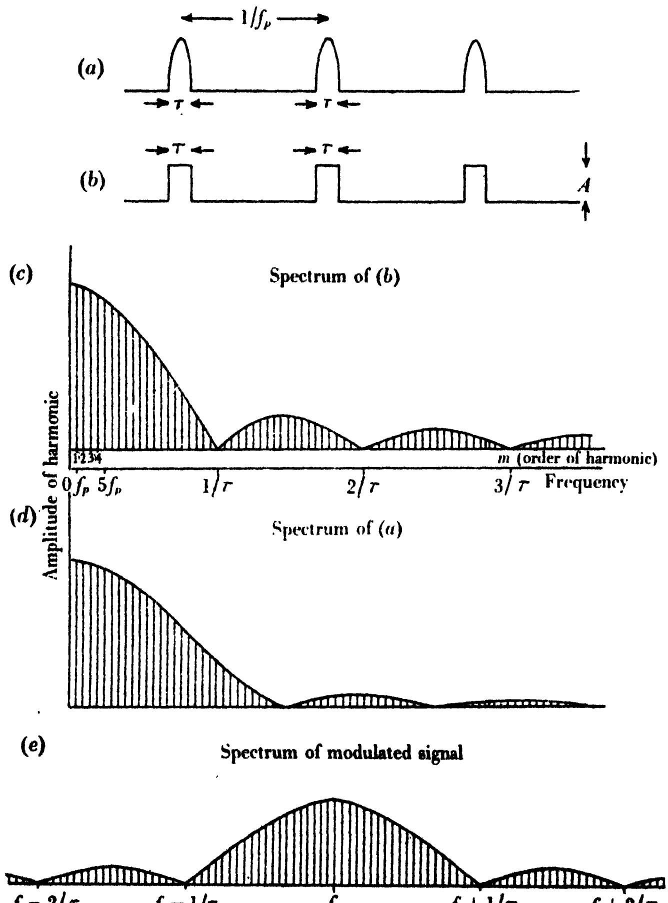


Fig. 2·2.

amplitude of the steady component is then $A\tau f_p$, and that of the mth harmonic $(2A/\pi m)\sin m\pi f_p\tau$. These results are shown in fig. 2·2*c*.* The envelope within which these harmonics lie is in fact

* We see that most of the energy lies at frequencies below $1/\tau$ ($m = 1/\tau f_p$) at which the intensity of the spectrum first falls to zero, and there are other zeros at $2/\tau$, $3/\tau$, etc.

the Fourier integral transform of a single pulse, depending only on τ and the pulse shape, being independent of f_p; this fact we shall use later. If the pulses, instead of being square-topped, are rounded, as in fig. 2·2*a*, the spectrum will be more like fig. 2·2*d*—the first zeros or minima coming at a higher frequency than before, but less of the energy will lie in harmonics above this zero or minimum.* For those familiar with directive array theory we may point out that these spectra are analogous to the main and side lobes of such an array; if the array 'filling' is tapered off towards the edge, the main lobe broadens slightly but the side lobes decrease in intensity.

Now when we use this function to modulate a carrier of frequency f_0, we know that each spectral component of the function (having frequency f_m, say) gives rise to two side bands having frequencies $f_0 \pm f_m$, so that we shall expect to obtain a spectrum similar to that shown in fig. 2·2*e*, in which most of the energy lies within the range $f_0 \pm 1/\tau$, the spectrum having actually discrete components spaced f_p apart. Now f_p is only of the order of 500 sec.$^{-1}$, while f_0 may be, perhaps, 5×10^8 sec.$^{-1}$, so that the separate components are very close and form an almost continuous spectrum. In fact, our theory has assumed pure amplitude modulation, in which the r.f. phases of successive pulses would be coherent (i.e. maintained in phase with an unmodulated continuous signal). In practice this is usually not the case—the oscillations build up anew for each pulse and the phases are quite random. This is equivalent to a frequency modulation of up to $\pm \frac{1}{2} f_p$ (just sufficient to change the phase $\pm 180°$ between successive pulses), which is just sufficient to blur the fine structure of the overall spectrum and convert the lines spaced f_p apart in fig. 2·2*e* into a continuous spectrum within the same envelope. Random irregularities in the pulse spacing have a similar effect.

We see, therefore, that the pass band required for a receiver is at the very minimum from $f_0 - 1/\tau$ to $f_0 + 1/\tau$, i.e. is $2/\tau$ wide. This will only pass components below the first zero of fig. 2·2*c*, and will distort the pulse shape considerably (note that reducing the components above the first zero tends to convert fig. 2·2*b* to fig. 2·2*a*, and removing them altogether will round the pulse off even further).

* If the pulses are semi-sinusoids, as fig. 2·2*a*, the Fourier envelope is given by the formula $\frac{4Af_p\tau}{\pi(n^2-1)}\cos\frac{n\pi}{2}$, where $n = 2\tau f$. This is the function plotted in fig. 2·2*d*, and has zeros at $n = 3, 5$, etc., compared with 2, 4, etc., for fig. 2·2*c*.

In order to transmit the pulse form faithfully a band width several times greater is required, say n/τ wide, where n may be 5 or so. When the pulse is steep rising, an n of 10 or more may be necessary to preserve this characteristic, which will be essential if very accurate range measurements are to be made. In fact, the accuracy possible for range measurements depends on this n—the accuracy of echo-delay timing being limited to $1/n$th of a pulse width, or τ/n.* This quantity n may be regarded as the order of the highest harmonic transmitted of the spectrum of a pulse, the 'fundamental' being the frequency for which the pulse length τ is just half a cycle. Although the spectrum is a continuous one and not a harmonic series, this concept gives a meaning to n which aids the memory—the pulses of fig. 2·2*a* being just half-cycles cut from a wave of the 'fundamental' frequency ($n = 1$).

As an example we see that for $\tau = 2{\cdot}5$ μsec. and $n = 5$ (i.e. extreme accuracy not required), a pass band 2 Mc./sec. wide is required, while for high accuracy with 1 μsec. pulses we might put $n = 10$ and use a 10 Mc./sec. wide pass band. The range accuracy in this case ($= \frac{1}{2}c\tau/n$) is ± 15 m.

2·4. Receiver characteristics—noise

In considering the characteristics necessary for a receiver to be used for pulsed radar, therefore, the first factor is the pass band required. Most receivers do not cut off suddenly at the ends of the band; usually instead they introduce phase distortion in a region of gradual cut-off, and attempts to sharpen the cut-off usually magnify the phase distortion. This is undesirable, since phase shifts of the higher frequency components of the pulse wave form tend to produce overshoots, i.e. pronounced oscillations on top of and following the pulse. A fairly gradual cut-off characteristic is therefore usually used, with say a 3 db. loss at $n = 5$ (i.e. at $f_0 \pm 2{\cdot}5/\tau$), or where particularly accurate ranging is required at, say, $n = 10$.

It is convenient to define a quantity called the 'energy band width' B to represent the effect of this gradual cut-off by an

* Note that the relation between range accuracy (δr) and band width (B) is the same as for the frequency-modulation system (cf. Chapter 1), viz.

$$\delta r = c\delta t/2 = c/2B,$$

since in our case $B = n/\tau$.

accurately defined figure. If G is the power gain of the receiver at any frequency f and G_0 that at mid-band frequency f_0, we take

$$B = \frac{1}{G_0}\int_0^\infty G\,df, \tag{2·1}$$

excluding any regions of unused (e.g. second channel) sensitivity. $n_b = B\tau$ will then usually be of the order of 5–10.

The other important constant of the receiver is sensitivity, which requires not only high gain but also low noise level. The problem of amplification, i.e. of producing sufficient gain, presents no difficulty at frequencies up to 500 Mc./sec., acorn types of valve being used early in the war, and later pressed-glass base types such as the Mullard EF 50. The latest development in this field is the use of grounded-grid amplifiers of which the S.T. and C. S25A and S26A are examples. At 3000 Mc./sec. and above, however, grave difficulties are encountered in obtaining radio-frequency amplification, and it is general practice to feed the aerial straight into a crystal mixer and perform all the amplification at an intermediate frequency. Superheterodyne reception is general for all radar frequencies. An i.f. of 45 Mc./sec. is common so that pass bands of 2–5 Mc./sec. width can be obtained without undue difficulty, although for Chain Home stations (cf. p. 115) an i.f. of 2 Mc./sec. was used. For special purposes i.f. amplifiers with band widths of 10–20 Mc./sec. have been constructed.*

In all cases, therefore, the limit of sensitivity is not amplification but noise level, and generally this is almost entirely the noise level originating in the first stage, be it r.f. amplifier or, as in the centimetre band, a mixer plus a first i.f. amplifier stage. Aerial noise is appreciable only on the very longest radar wave-lengths, so it is most important to reduce receiver noise to a minimum. There is an irreducible minimum of noise associated with any resistor—this is the thermal or Johnson noise which arises from the random thermal motions of the electrons in the resistor, and whose mean square value is given by

$$\overline{V^2} = 4kTR\,df, \tag{2·2}$$

where k is Boltzmann's constant, T the absolute temperature, R the value of the resistance concerned, and only the noise-voltage components within a frequency band of width df are considered. In any

* See, e.g., Lewis, W. B., *J. Instn Elect. Engrs*, vol. 93, part III A, p. 272 (1946).

receiver or amplifier noise originates in addition in the valves. The Shot effect is noise due to the random arrival of electrons at the anode in what is normally regarded as a steady current; its magnitude depends on space-charge conditions (e.g. on whether the current is space-charge limited), and its detailed discussion is outside the scope of this book. In tetrodes and pentodes 'partition noise' may also arise. These matters and the design of low noise valves are treated in a companion volume in this series.

2·4·1. *Noise factor*

A quantity called the noise factor (N) has been defined for any receiver as a measure of how much worse it is than the theoretical minimum set by the Johnson effect, and is given by

$$N = \frac{N_o S_i}{N_i S_o}, \tag{2·3}$$

where N_o = noise power available at the receiver output,

S_o = signal power available at the receiver output,

S_i = signal power available from the input source, and

N_i = noise power within the 'energy band width' of the receiver available from a resistance or linear network having the same impedance as the input source and giving Johnson noise corresponding to a standard temperature of $T_0 = 290°$ K.

The word 'available' is used in case the receiver input is not matched to the input source (e.g. aerial feeder) impedance; the power available into the optimum load is to be taken in these expressions. The noise factor is a pure number, but is sometimes expressed in decibels. If the receiver pass band is sufficient to pass the signal wave form substantially unchanged, S_o/S_i is just the power gain of the receiver (less any loss due to an input mis-match), and N, as its name implies, depends only on noise levels. We may define a quantity N_e as the effective noise power at the input of the receiver such that N_o/N_e is also the power gain of the amplifier, and we then see that $N_e = NN_i$. Since in a well-designed amplifier very little contribution to the total noise arises from stages after the first, the concept N_e as noise in watts at the input terminals expresses quite well the amount of noise originating in this first stage.

The noise factor as defined above depends on the input matching arrangements (cf. fig. 2·3). The optimum power match (i.e. condition for maximum power transfer) occurs, as is well known, when R_s, the impedance of the input source, is made equal to R_i, the receiver input impedance.* When this is so, N_i is readily calculated; the noise voltage $\overline{V}$ in the source produces a current $\overline{V}/2R_s$ in the load, so that the available power is $i^2R_i = \overline{V}^2/4R_s$ (since $R_i = R_s$). Using equation (2·2) we see that

$$N_i = \overline{V}^2/4R_s = kT_0B. \tag{2·4}$$

This is true in all cases, since N_i is defined as the available power, although only at the optimum power match is it the actual noise power fed into the receiver from a resistance source. It must be

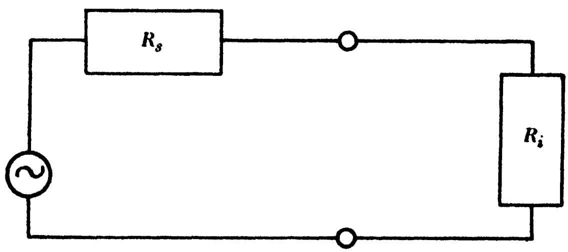


Fig. 2·3.

realized, however, that the receiver input impedance R_i also produces Johnson noise, even in a theoretically perfect receiver, so that $N_e \geqslant 2N_i$ providing $R_s = R_i$. With optimum power match, therefore, N can never be less than 2. Lower values may, however, be attained if R_i is made greater than R_s, and the ultimate limit in this case is $N = 1$.

It therefore may happen that the optimum match for best signal-to-noise ratio may not be the same as the optimum power match. If, as at long wave-lengths, aerial noise is greater than receiver noise, the optimum power match will also be nearly optimum for signal-to-noise ratio, although the maximum will be flat. Also, if the noise factor of a receiver is high, as it is, for example, in the centimetre band, the noise originating in R_s will be negligible and

* For simplicity in this treatment we assume the impedances to be non-reactive. The optimum power match in the general case is when Z_s and Z_i are complex conjugates.

again optimum power match will be desired—to obtain as much signal as possible for a fixed valve noise. But when aerial noise is low as well as the receiver noise factor, it pays to depart from the optimum power match so as to improve the noise factor; any loss of signal strength can easily be made up by increasing the gain. Any such adjustment is best made empirically, but it is important to realize that the noise factor may vary with input matching.

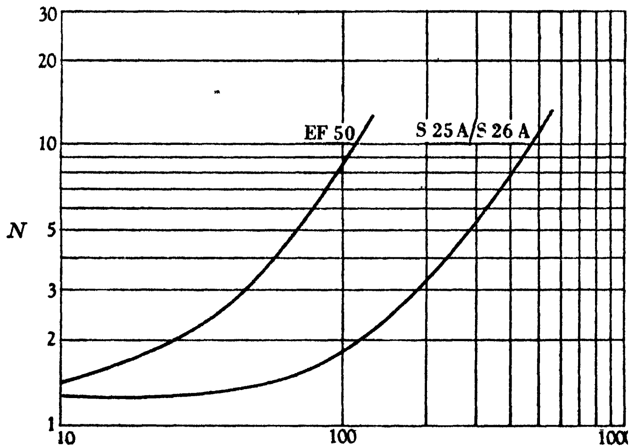


Fig. 2·4. Noise factors for typical r.f. amplifier input stages.

The matter is further complicated because the equivalent temperature of the aerial may not be anything like 290° K. This matter is dealt with more fully in § 3·3, but it may be stated generally that above 3 m. wave-length aerial noise is generally predominant over receiver noise, whereas for shorter wave-lengths the reverse is true. In the former case the noise factor has to be modified; we may divide $N_e = NkT_0B$ into $(N-1)kT_0B$ originating in the receiver and $kT'B$ originating in the aerial, T' being the effective aerial temperature (substituted for T_0 in this term), so obtaining an effective noise factor $N-1+T'/T_0$ instead of the original N.

With these reservations that actual noise factors depend somewhat on conditions, especially when N does not greatly exceed unity, we give some comparative curves (fig. 2·4) which indicate approximately what values of noise factor are attainable at various wave-lengths. For the crystal mixer inputs used in the centimetre

band it is even more difficult to give reliable figures for noise factors. Lewis gives the expression $N = (N_{\text{i.f.}} + T_c/T_0 - 1)/G_m$ for the over-all noise factor, where $N_{\text{i.f.}}$ is the noise factor from the i.f. input onwards, T_c the noise temperature of the crystal, and G_m the conversion gain of the mixer (in terms of 'available' power). T_c varies from about T_0 to $4T_0$, $2T_0$ being an average figure, and G_m is stated to be from -6 to -9 db. at 3000 Mc./sec. We may therefore estimate that the over-all noise factor at 3000 Mc./sec. will be somewhere near 20 (13 db.).

2·5. Perception factor—limiting noise sensitivity

In determining the minimum detectable signal there is an additional factor attributable not to the receiver proper but to the video circuits and display. This quantity is called the 'Perception Factor' (Π), and is defined as the ratio of signal power* to noise power at the second detector corresponding to the smallest signal which is detectable at the display. The values of Π attained in practice lie generally between 1 and 4; if they exceed the latter figure it is usually because the video circuits are badly designed or the display controls badly adjusted. Perception factors less than unity are not theoretically impossible, but can usually only be obtained with circuits integrating the responses from several successive traces and so smoothing out the noise. In practice a value of $\Pi = 1 \cdot 5$ is considered a very good performance.

We are now in a position to define the 'limiting noise sensitivity', which is the minimum power available at the receiver which will give a detectable signal. We shall denote this by the symbol P_m. Both Π and P_m, of course, depend markedly on the type of display used and on the pulse shape and duration. From the definitions of the quantities concerned it is readily seen that

$$P_m = \Pi N N_i = \Pi N k T_0 B.$$

The value of this quantity will be required in the following chapter.

2·6. Two numerical examples

We conclude this chapter with two numerical calculations of the limiting noise sensitivity. For one case we consider a 10 cm. wavelength receiver using a crystal mixer with a noise factor N of

* Usually the power at the peak of the pulse (cf. page 25).

20 (13 db.). We may require a high range-accuracy (cf. § 2·3) using a 1 μsec. pulse and $B = 10$ Mc./sec., and let us take $\Pi = 1{\cdot}5$ (a good average value). Then, since $k = 1{\cdot}37 \times 10^{-23}$ joule/° K.,

$$P_m = 1{\cdot}5 \times 20 \times 1{\cdot}37 \times 10^{-23} \times 290 \times 10^7$$
$$= 1{\cdot}19 \times 10^{-12}\,\text{W}.$$

Since $P = V^2/R$, this represents a voltage across a matched 75-ohm cable of $\sqrt{(75 \times 1{\cdot}19 \times 10^{-12})} = 9{\cdot}4 \times 10^{-6}$ V.

For the more favourable case of a 200 Mc./sec. receiver using an S26A grounded-grid triode input, $N = 3{\cdot}4$, while if the band width is 4 Mc./sec., and Π still $1{\cdot}5$, we have

$$P_m = 1{\cdot}5 \times 3{\cdot}4 \times 1{\cdot}37 \times 10^{-23} \times 290 \times 4 \times 10^6$$
$$= 8{\cdot}1 \times 10^{-14}\,\text{W.},$$

representing 2·46 μV across 75 ohms.

We see thus that the order of magnitude of P_m for the cases likely to be met with in radar is from 1 to 0·1 $\mu\mu$W., corresponding to some 2–10 μV. across a 75-ohm cable.

Chapter 3

FACTORS AFFECTING THE PERFORMANCE OF A RADAR EQUIPMENT

In the previous chapter we have considered the two most obvious factors which affect the performance of a radar equipment, viz. the power P_s of the pulse sender, and the minimum power P_m to which the receiver will respond, a quantity which we have called the 'limiting noise sensitivity'. In the present chapter we examine the other factors which determine the performance, using that word primarily in the sense of maximum range, but also to include considerations of how the maximum range varies with direction (e.g. with angle of elevation from a ground station), and whether the form of any such variation is suited to the purpose for which a given radar equipment was designed. The factors with which we shall be concerned are, principally, the nature of the reflecting target, the gain and directivity of the arrays, and the conditions of propagation. Finally, we consider the problem of the choice of a suitable wave-length for the radar in the light of these factors.

We start, therefore, from the sender power and the receiver sensitivity, or rather, from their ratio P_s/P_m. It is usual to express P_s as the power at the peak of the pulse, but if the mean power taken over the pulse duration τ is used, it is immaterial, *provided* that P_m is expressed in the same way. P_m (and the perception factor Π) can only be defined for a given pulse duration and pulse shape, so that no inconsistency can arise whatever basis of reckoning is used.

3·1. The gain of directive aerials

In order to obtain sufficient maximum range, the use of directive aerials concentrating the radiation in a desired direction is essential for practically all radar purposes. This is particularly true at the shorter wave-lengths, where highly directional aerials are the rule, whereas at 12 m. a few dipoles suitably phased suffice to give the directivity necessary. An important quantity is therefore the 'gain'

of the aerial, a factor which expresses the increase in the power radiated in a given direction over that which would be radiated if the same total power were fed into a completely non-directional or 'isotropic' aerial. It should be noticed that the gain is a function of direction, and must necessarily be less than unity in some directions to make up for the concentration of energy in others; sometimes, however, 'the gain' of an aerial is referred to, meaning the gain measured in the direction of maximum radiation.

Our definition of gain, in terms of the isotropic aerial as unity, is not universally accepted, although it is the most fundamental. An earlier standard, which is still commonly used, is to take the gain of a half-wave aerial (in its equatorial plane) as the unit. This is advantageous in experimental work, since the half-wave aerial is easily realizable, while the isotropic radiator is not. Indeed, no simple aerial can be isotropic; electromagnetic radiation cannot be longitudinally polarized, and all simple types of aerial have a direction along which no radiation occurs. But the advantages in computations of using the isotropic aerial as the unit of gain are such that we propose to use it throughout. When it is necessary to refer to gains measured in terms of the half-wave aerial unit, we shall use the symbol G' instead of G. Since the gain of a half-wave aerial is 1·64, we have the universal relation $G = 1{\cdot}64G'$. A third possible unit of gain, the equatorial gain of a Hertzian dipole (1·5 times our unit), we shall not employ.

We have defined aerial gain in terms suitable for a radiating aerial, but it is a well-known example of the theorem of reciprocity that the gain as a receiving aerial is the same figure, the radiation being supposed incident along the same direction as that for which the gain as a radiator was defined, and the polarization of the radiation being unchanged. The actual definition of gain for the receiving case is as a measure of the amount of energy absorbed by the aerial from a plane wave incident in the desired direction. The related concept of effective absorption area (A_a) of an aerial is also useful; this is defined by saying that the power 'available' from such an aerial is just equal to the energy flux across an area A_a placed normal to the incident radiation. The term 'available', as in § 2·4·1, is used to mean that which is obtained in the load when the aerial is correctly loaded for maximum power absorption. The gain of a receiving aerial is then by definition proportional to A_a, and since the effective

absorption area of our unit gain aerial (the isotropic aerial) is $\lambda^2/4\pi$,* we obtain the general result that

$$A_a = G\lambda^2/4\pi. \tag{3·1}$$

It may be noted, where gains with respect to a half-wave aerial are used, that since $G = 1{\cdot}64G'$, A_a is also approximately equal to $G'\lambda^2/8$,

$$A_a = G'\lambda^2/8. \tag{3·1 a}$$

Since this concept of energy flux will be in frequent use, a few words on the subject are desirable. In electromagnetic theory it is represented by the Poynting vector $c[\mathbf{E}\wedge\mathbf{H}]$, it being implicit in the theory that for a wave travelling in free space E in e.s.u. and H in e.m.u. are numerically equal. Now H is related to a displacement current which can be considered as flowing in a sheet parallel to E in any wave front, and changing our units we find that E in volts per metre and the displacement current i in amperes per metre are in phase and have a fixed ratio equal to

$$4\pi c \times 10^{-9} = 120\pi = 377 \text{ ohms}.$$

This figure may be regarded as the impedance of free space; if we could remove the space beyond it, all radiation would be absorbed by a thin film having a resistivity of 120π ohms per square (a square of any size having the same resistance between opposite sides), so that such a film could be said to be the correct termination for free space. It follows immediately that the energy flux in a radiation field of E V./m. is just $E^2/120\pi$ W./sq.m.; we call this quantity F and write

$$F = E^2/120\pi = E^2/377. \tag{3·2}$$

We may therefore deduce the power available from an aerial of gain G placed in a radiation field of strength E V./m. as

$$P_a = \frac{G\lambda^2}{4\pi} \times \frac{E^2}{120\pi} = \frac{G\lambda^2E^2}{480\pi^2}. \tag{3·3}$$

The actual power transferred to the receiver and the voltage appearing at the receiver terminals depend on the matching conditions, but if the termination is 'correct' (i.e. optimum power

* Cf. Appendix 2.

match, cf. §2·4·1), the input voltage V and the shunt resistive component R of the input impedance are related by $P_a = V^2/R$, so that

$$V = \sqrt{\left(\frac{GR}{30}\right)}\frac{\lambda E}{4\pi}.$$

In long-wave practice it is common to talk of an effective aerial height H defined by $V = HE$; while we can do this for radar frequencies, it is a less useful concept, since H depends very markedly on the matching arrangement. Indeed, for radar we prefer to work throughout in terms of power, avoiding as far as possible the concepts of field strength and voltage. Thus we calculate the energy flux at a distance r from an aerial radiating power P_s and having gain G in the direction concerned, as being G times as great as a total power P_s distributed uniformly over the surface $4\pi r^2$ of the sphere of radius r surrounding the source, thus obtaining

$$F = \frac{P_s G}{4\pi r^2}, \tag{3·4}$$

and continuing our calculations in the same way without ever using the concept of field strength. This simple method follows directly from our adoption of the isotropic aerial as the unit of gain; examples of the method will be given later in this chapter.

We have seen that a high-gain aerial necessarily has a high effective absorption area; the physical area of the array cannot be less than this quantity, so that we see that a large area array is essential to the production of high gain. It is clear that a radiating array can only have a high gain in one direction by concentrating most of the radiation in a narrow beam, which must mean an array of large aperture. This is an elementary result of diffraction theory, viz. that to concentrate most of the energy in a beam of angular width $1/n$th of a radian an array of linear dimensions of the order of $n\lambda$ is required. Thus consideration of the array either as a radiator or as a receiver of energy leads to the same result, that a large area array is essential to high gain, the area required being also proportional to λ^2.

Consider, for example, the broadside array (fig. 3·1), formed of half-wave aerials all fed in phase and forming a rectangular array of aerials spaced $\frac{1}{2}\lambda$ in either direction. If there are n such 'dipoles', the area of the array (taken from $\frac{1}{4}\lambda$ above the top row of dipoles

to $\frac{1}{4}\lambda$ below the bottom row) is $n\lambda^2/4$. On the line of shoot (normal to the array) the radiated field is n times, and the energy flux n^2 times, that due to the same current in one dipole, whereas the power needed to drive the array is only n times that for a single dipole, assuming the radiation resistance of the dipoles unchanged. Approximately, therefore, the gain of such an array is n times that of a half-wave aerial, i.e. $G' = n$ so that the effective area A_a is $G'\lambda^2/8 = n\lambda^2/8$ in our case, or half the actual area of the array. If, however, we use a suitable aperiodic reflector behind the dipoles

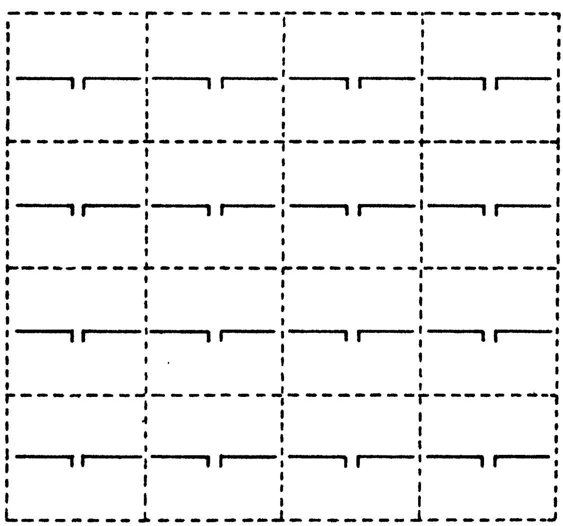


Fig. 3·1. Broadside array. Each dotted square $= \frac{1}{2}\lambda \times \frac{1}{2}\lambda$ (area $\frac{1}{4}\lambda^2$) contains one $\frac{1}{2}\lambda$ aerial. The whole aerial is $2\lambda \times 2\lambda = 4\lambda^2$, $n = 16$.

to suppress the radiation behind the array, we approximately double the array gain, and such an array has an effective absorption area substantially as large as the array itself.

For wave-lengths below about 1 m., however, broadside aerials are little used, mainly because high gains are required and the phasing of a large number of separate aerials presents difficulties. The paraboloidal aerial (fig. 3·2) is much used at shorter wave-lengths. In this, which is quasi-optical, an aperiodic reflector converts the divergent radiation from a dipole into a plane wave front. As with a broadside array, the beam has a finite width due to diffraction. With the usual arrangement the intensity of the plane

wave front produced is greatest at the centre of the aperture and falls off towards the edges. This reduces the gain somewhat, and a common figure for an aerial of circular aperture is that A_a is about three-quarters of the physical area of the aperture.

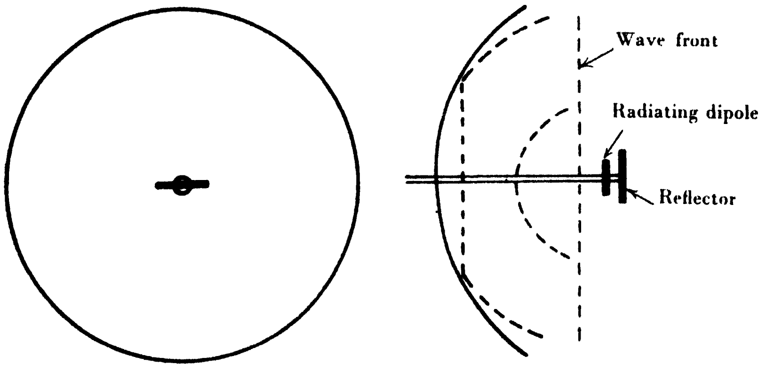


Fig. 3·2. Paraboloidal aerial.

3·1·1. *Polar diagrams of directive aerials*

It is clearly desirable to use high-gain aerials to obtain long ranges with radar, but there are limits to this process. One limit is that the area of the array may become prohibitive—another may be that the beam becomes too narrow for the purpose in view. For many applications a pencil beam is suitable, gain being obtained by beaming in both angular coordinates, but sometimes a 'fan-beam' is necessary, in which beaming in only one coordinate is allowable, e.g. it may be possible to beam in azimuth but at the same time be necessary to illuminate targets at all elevations. The gain obtainable is generally less in the latter case.

Further, it is not possible using a finite array to concentrate all the energy uniformly within the beam width; all actual beams are strongest in the centre, the intensity falling off gradually towards the edges, while frequently unwanted minor lobes appear outside the main beam. Even though these do not represent any great loss of power, they may be a nuisance operationally in giving rise to unwanted echoes. Consider, for example, a rectangular aperture in the x-y plane extending from $x = -\frac{1}{2}a$ to $x = +\frac{1}{2}a$ and from $y = -\frac{1}{2}b$ to $y = +\frac{1}{2}b$ uniformly filled with radiating current elements all in phase (cf. fig. 3·3*a*). This will radiate a beam along OZ and

OZ', and the beam width is readily calculated, being a standard result available in any text-book of Physical Optics, as the diffraction pattern of a rectangular aperture. Putting for convenience $k = 2\pi/\lambda$,

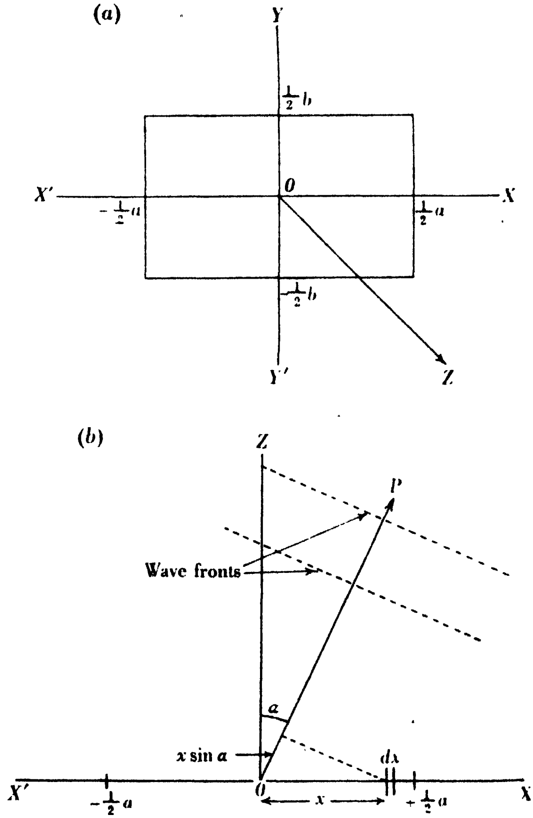


Fig. 3·3.

the diffraction intensity for a direction making an angle $\frac{1}{2}\pi - \alpha$ with OX and $\frac{1}{2}\pi - \beta$ with OY is proportional to

$$\frac{\sin(\frac{1}{2}ka\sin\alpha)\sin(\frac{1}{2}kb\sin\beta)}{ka\sin\alpha\ kb\sin\beta}.$$

The half of this expression which is a function of α is readily obtained from a consideration of fig. 3·3 b, which is a section in the OXZ plane. The radiation from a strip lying between x and $x+dx$ has intensity proportional to dx and has to travel a distance $x\sin\alpha$

less than the radiation from O to reach any wave front corresponding to propagation along OP. With the usual convention we may therefore write the intensity as $e^{jkx\sin\alpha}\,dx$. The total radiation vector is therefore

$$\int_{-\frac{1}{2}a}^{\frac{1}{2}a} e^{jkx\sin\alpha}\,dx$$

$$= \frac{1}{jk\sin\alpha}\left(e^{\frac{1}{2}jka\sin\alpha} - e^{-\frac{1}{2}jka\sin\alpha}\right)$$

$$= 2\sin\left(\tfrac{1}{2}ka\sin\alpha\right)/k\sin\alpha,$$

which apart from a constant factor $2a$ is the expression required. Note that the result is real, i.e. the resultant has always the same

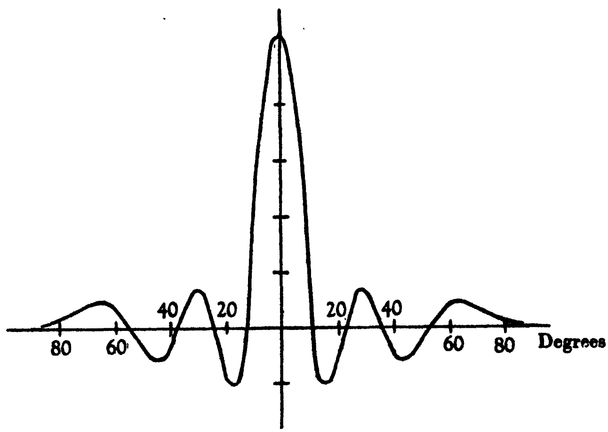


Fig. 3·4. Field strength pattern for array ($a = 5\lambda$).

phase (apart from a possible -1 factor) as the wavelet originating from O, the centre of the array. This result is general for all symmetrical arrays, whether or not the radiating elements are uniform across the whole aperture.

Fig. 3·4 shows this function plotted for the case of $a = 5\lambda$. It represents signal strength, and must of course be squared to represent power density. It will be seen that most of the energy lies between the first two zeros, which lie at $\alpha = \pm\sin^{-1}(\lambda/a)$ (and in the other plane $\beta = \pm\sin^{-1}(\lambda/b)$). The beam width is therefore inversely proportional to the aerial aperture in the same direction. It can be shown that if the illumination across the aperture tapers towards the edges the beam will be somewhat wider, but the side

lobes will be smaller—the result is similar to that shown in fig. 2·2 for the Fourier transforms of the square-topped and semi-sinusoid wave forms. For a circular aperture uniformly filled and of diameter d, the beam width to the first zero is $\pm\sin^{-1}(1{\cdot}22\lambda/d)$, and the side lobes are smaller than for the rectangular apertures. For the paraboloidal aerials of circular aperture commonly used for radar this result holds approximately in a plane parallel to the **H**-vector, while in the other plane the illumination across the aperture is less uniform, due to the directivity of the feeding dipole, and the beam is even wider and the side lobes smaller.

3·1·2. *The effect of reflexions from the ground*

A major consideration for many radar ground and ship equipments is the effect on the vertical polar diagram and gain of the existence of reflected radiation from the land or sea in front of the aerials. For one thing, this radiation makes it impossible to obtain good field strengths at low angles of elevation without using aerials many wave-lengths above the earth's surface—a severe limitation at long wave-lengths.

For grazing incidence the reflexion coefficients of both land and water are -1, but for rays making other angles (α) with the surface the complex reflexion coefficient (ρ) depends on both the angle and the frequency. Sea water behaves to radio waves approximately as a medium of dielectric constant $K = 80$ and conductivity $= 10^{10}$ e.s.u.,* so that its equivalent complex refractive index is given by $\mu^2 = K - 2j(\sigma/f)$, from which the reflexion coefficient can be deduced using the Fresnel equations

$$\rho = \frac{\sin\alpha - \sqrt{(\mu^2 - \cos^2\alpha)}}{\sin\alpha + \sqrt{(\mu^2 - \cos^2\alpha)}} \tag{3·5}$$

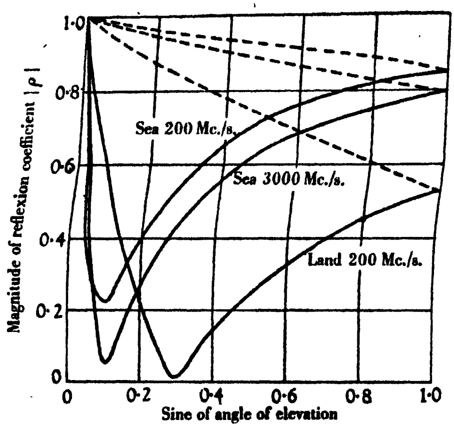
for horizontal polarization, and

$$\rho = \frac{\mu^2\sin\alpha - \sqrt{(\mu^2 - \cos^2\alpha)}}{\mu^2\sin\alpha + \sqrt{(\mu^2 - \cos^2\alpha)}} \tag{3·6}$$

for vertical polarization.

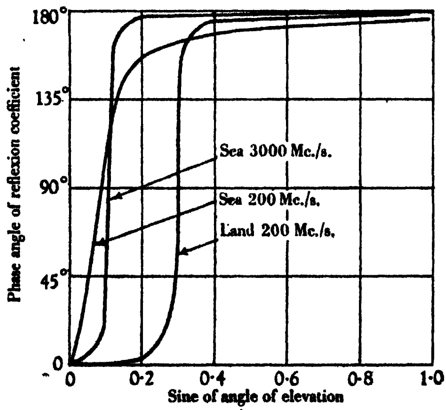
Land conditions vary somewhat with the nature and wetness of the soil, but average values* are $K = 10$ and $\sigma = 5\times10^7$ e.s.u. From these values the reflexion coefficients shown in fig. 3·5 are deduced. It will be noticed that for $\alpha = 0$ both these expressions

* Burrows, C. R., *Bell Syst. Tech. J.* vol. 16, p. 45 (1937).

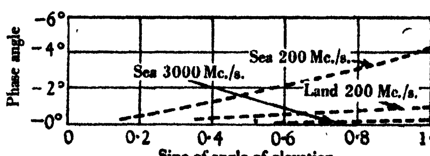


(*a*) Magnitude of reflexion coefficient

——— Vertical polarization ----- Horizontal polarization



(*b*) Phase of reflexion coefficient (vertical polarization)



(*c*) Phase of reflexion coefficient (horizontal polarization)

Fig. 3·5. Complex reflexion coefficient calculated for land ($K=10$, $\sigma=5\times10^7$ e.s.u.) and sea ($K=80$, $\sigma=10^{10}$ e.s.u.). Note the effect of raising frequency from 200 to 3000 Mc./sec. for the case of sea—at the higher frequency the phase passes rapidly through 90° while the magnitude is small. For land this is already the case at 200 Mc./sec., the reflexion coefficient for all higher frequencies being substantially identical with that shown. (Phase angle $=0°$ for $\rho=-1$.)

reduce to $\rho = -1$, which approximate value is often used for calculations at low angles of elevation, particularly for horizontal polarization, for which (cf. fig. 3·5) the approximation remains fairly good at moderate elevations. At sufficiently high frequencies, roughly above 100 Mc./sec. for land and 2500 Mc./sec. for sea water, the term in σ becomes negligible, and the simple formula $\mu = \sqrt{K}$ can be used; in the case of land this result applies for most of the radar frequency-band, and for sea it applies for all centimetre waves. In this region the reflexion coefficients become independent of frequency and are always real, i.e. no change of phase (other than 180°) occurs on reflexion.

A particular case of some importance is when the terrain in front of the aerial is flat. This occurs at sea, and from cliff coastal sites, and inland sites are often chosen so that it shall be true to a sufficient degree of approximation also. It may be stated at once that, following Rayleigh, we may treat as substantially flat any surface which has small bumps and dips departing not more than $\pm\lambda/(8\sin\alpha)$ from the median plane, for radiation of wave-length λ and grazing angle α. 'Small' generally means small compared with the area from which the energy would be reflected on ray theory, or small compared with a Fresnel half-wave zone.

For such a flat site, and taking $\rho = -1$ initially, we may consider in detail the case of a symmetrical aerial whose centre is at height h above the site (fig. 3·6). By 'symmetrical' we imply that the radiation at elevation α and at $-\alpha$ are equal in magnitude and in phase, referred to the ray emanating from the centre A of the aerial. Let O be the foot of the perpendicular from A on to the site, and A' the image of A. Then if P be the 'bounce-point' of the reflected ray, and N the foot of the normal from A on to $A'P$, we have a resultant radiation at elevation α made up of two equal-intensity rays with a path difference of $A'N = 2h\sin\alpha$, one having suffered a change of phase of π on reflexion. Their phase difference is therefore $\pi + 2kh\sin\alpha$, where $k = 2\pi/\lambda$ as before, and if the magnitude of either is $f(\alpha)$—the free-space directivity of the array—the magnitude of the resultant will be $2f(\alpha)\sin(kh\sin\alpha)$, as will be seen from the vector diagram, fig. 3·6*a*. The phase of the resultant is also seen to be in quadrature with the reference phase (that of a wavelet from O). The resulting magnitude is plotted in fig. 3·7 for two values of h. We see that alternate lobes and gaps are produced in the vertical

radiation pattern, and that in the maximum of the lobe the field strength is twice what it would be in the absence of reflexions from the site. This corresponds to multiplying the gain fourfold. On the other hand, the gaps produced at other angles of elevation are a serious disadvantage. It is also seen that for a high aerial the lobes

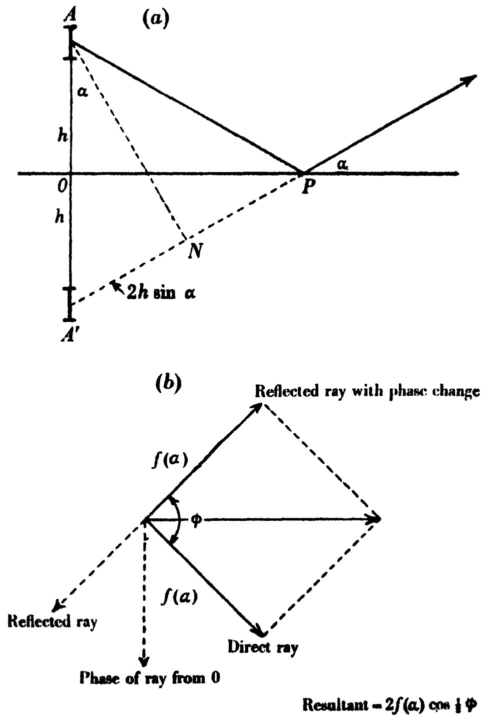


Fig. 3·6. Illustrating effect of ground reflexions.

and gaps are close together, while for a low aerial they are widely spaced, but that the latter arrangement gives a low field strength for low angles of elevation.

We give in fig. 3·8 an illustration of the effect of the departure of ρ from -1, taking the case (both polarizations) of land of $K = 10$ and such a high frequency that σ can be neglected. We see that the imperfect reflexion produces a measure of gap-filling, at the expense however of gain at the lobe maximum. Had the σ term been appreciable, producing a phase shift on reflexion, the position of the

minima would also have been different. There are other methods of gap-filling, such as tilting the array so that the reflected ray is reduced in intensity, but the matter is too specialized to pursue in detail here. It should be clear, however, that for pencil beam and

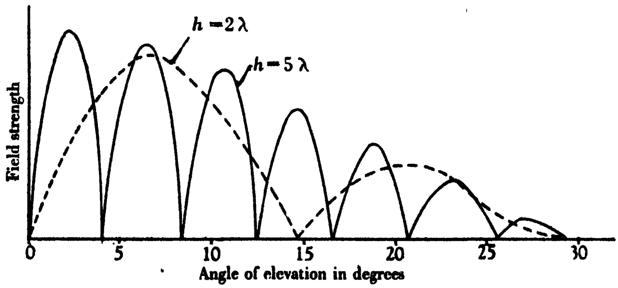


Fig. 3·7. Vertical polar diagram of an array of vertical aperture 2λ at two different heights. $\rho = -1$.

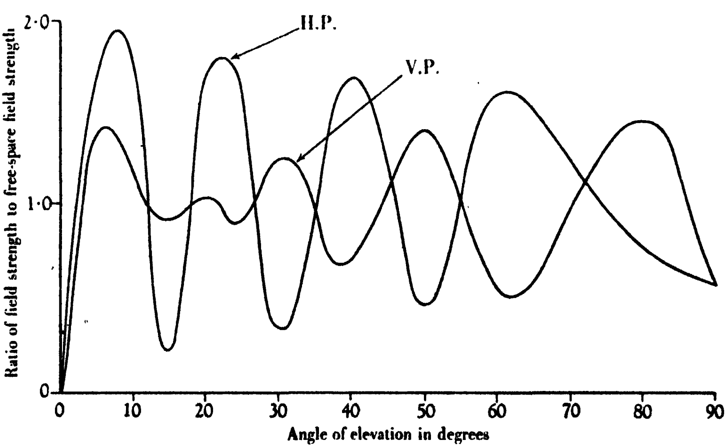


Fig. 3·8. Curves illustrating effects of imperfect reflexion. $h = 2\lambda$.

similar equipment for which the beam is tilted up so as to be clear of the horizon, there will be no substantial effects due to earth reflexions.

3·2. Equivalent echoing area of a target

We turn now to the next factor, the target which reflects the radar waves. This is merely a scattering agent, and we may measure its efficiency as such by the equivalent scattering area A_s, defined on

the lines of A_a, such that the total scattered energy is that intercepted by an area A_s placed normal to the incident ray. We are not strictly interested, however, in the total scattered energy, but in the energy scattered back along the direction of incidence. The reradiation of the target may not be isotropic, but if it has a gain of G back along the direction of incidence, we may define a new area A, called the echoing area, such that $A = A_s G$. Then A is such that the reflected signal is that which would be produced if the power in the incident wave intercepted by area A were reradiated by an isotropic radiator at the target. The echoing area A is fundamental to all radar problems, hence we write the symbol without a subscript.

There is an alternative concept also used for measuring echoing power; this has the dimensions of a length L, called the echoing length, and is defined by $E_e/E_i = L/r$, where E_i is the incident field strength at the target, and E_e is the echoed field strength at distance r back along the direction of incidence. It may readily be proved that the two concepts are related by the equation

$$A = 4\pi L^2. \tag{3·7}$$

For certain simple cases echoing areas may be calculated directly. Such cases include a resonant half-wave aerial, a sphere, and so on. These calculations are given in Appendix 2. For other targets of complex form, such as an aircraft, we can only obtain the echoing areas empirically, although in one interesting case, the German V2 rocket, quite good estimates were made of the echoing areas to be expected for different wave-lengths and aspect angles by calculation before any measurements were possible. We shall consider the information available about aircraft targets in the following section, following this with a consideration of surface targets (ships, submarines, etc.), for which rays reflected off the sea or the earth's surface complicate matters somewhat.

3·2·1. *Echoing areas of aircraft*

Many flights have been made during the war to test the performance of various radar stations, and much evidence has accumulated as to the order of magnitude of A and L for aircraft of different sizes and for wave-lengths from about 12 m. down to 3 cm. Most of these flights were radial from and back to the stations concerned, so that most of the information concerns head and tail aspect, and the figure usually given is for the mean of these two aspects. There are

quite complicated variations with aspect, but in general it may be said that for all wave-lengths appreciably shorter than twice the wing span of the machine, the average echoing constants are substantially independent of wave-length, for horizontal polarization at least. For vertical polarization this statement should perhaps be restricted to wave-lengths less than twice the vertical dimension of the aircraft, but here less information is available at the long wave-lengths. This is roughly what would be expected from an object of complex form. We may therefore give in table 1 a few figures which may be taken as typical, with the caution that for some target aspects L may fall as low as 10 % or A as 1 % of the figure given, while for other conditions the figures given may be considerably exceeded.

Table 1. *Average echoing constants*

Type of aircraft	Wing span (ft.)	L (m.)	A (sq.m.)
Hurricane	40	0·65	5·4
Beaufighter	58	1·0	13
Ju 52	96	1·7	36

As far as aspect variations are concerned, it has frequently been observed that large signals are obtained from a test aircraft on the turn at a range such that it is visible with neither head nor tail aspect. Very little consistent information has been obtained in the field on aspect variations, but a certain amount of detailed measurement has been made on models, though the results are too complicated to summarize. At the shorter wave-lengths rapid changes of the echoed signal are noticeable, and are attributed to vibrations of the aircraft structure, causing the wavelets scattered by different parts of the aircraft to come rapidly into and out of phase with one another. A further complication arises due to the rotating airscrew; this gives a modulated reflected wavelet causing modulation of the response of the whole aircraft having strong Fourier components at frequencies corresponding to multiples of the airscrew rotation frequency. Several harmonics of this frequency usually occur, some of them related to the number of blades on the airscrew, and for multi-engined planes the pattern may become very complicated. This 'propellor modulation' has been observed by photographing the amplitudes of successive pulses received from a given aircraft, and may occasionally be of use in making certain that a given target is not jet-propelled. At certain aspects this propellor modulation

may not appear, so that it might be difficult to be sure that an aircraft was jet-driven.

It should also be mentioned that crossed polarization is sometimes used, the radar sending aerial being horizontally polarized and the receiver vertically, or vice versa (cf. § 8·0). The polarization-conversion echoing areas for aircraft targets are somewhat smaller than the ordinary echoing areas, on the average by a factor of four perhaps, but are still quite large. They tend, too, to be low for symmetrical aspects such as head-on and tail-on, as might be expected, but for some oblique aspects may even be larger than the ordinary echoing area.

3·2·2. *Echoing constants of surface targets*

For the detection of ships, submarines, etc., on the sea, or tanks, etc., on land, matters are somewhat complicated by the interference between direct and reflected waves. Little information is available about echoing areas of targets on land—on land there tends to be so much 'clutter' in the form of reflexions from terrain irregularities that the subject has not been studied very seriously. We shall therefore restrict our statements to targets at sea, the sea being supposed smooth enough to give specular reflexion (cf. § 3·1·2 for the Rayleigh criterion).

For aircraft detection, as we have seen, two cases arise, according as to whether reflexions from the ground are or are not important, and in the former case the matter is taken care of by regarding these reflexions as altering the gain of the radar aerial by a factor depending on the elevation angle. This is possible because the reflexion takes place near the radar (fig. 3·9*a*) and the target is in an area of substantially constant field strength. But for surface-vessel detection the reflexion takes place near the vessel (whether the radar is on a cliff or an aircraft is immaterial), with the result that the echoing target is not placed in a constant incident field, but in one which varies across the surface of the target, being zero* at the water-line. The concept of echoing area in its original form is therefore not applicable.

Two cases arise, depending on the height (H) of the vessel and on the angle of elevation of the radar (α) as seen from the target.

* We assume $\rho = -1$, which is usually very nearly true for the cases met with in surface-vessel detection.

If $2H \sin \alpha$ ($A'N$, fig. 3·9c, cf. fig. 3·6) $< \frac{1}{2}\lambda$, the top of the vessel lies below the maximum of the first interference lobe, and the incident (resultant) field strength increases monotonically from the

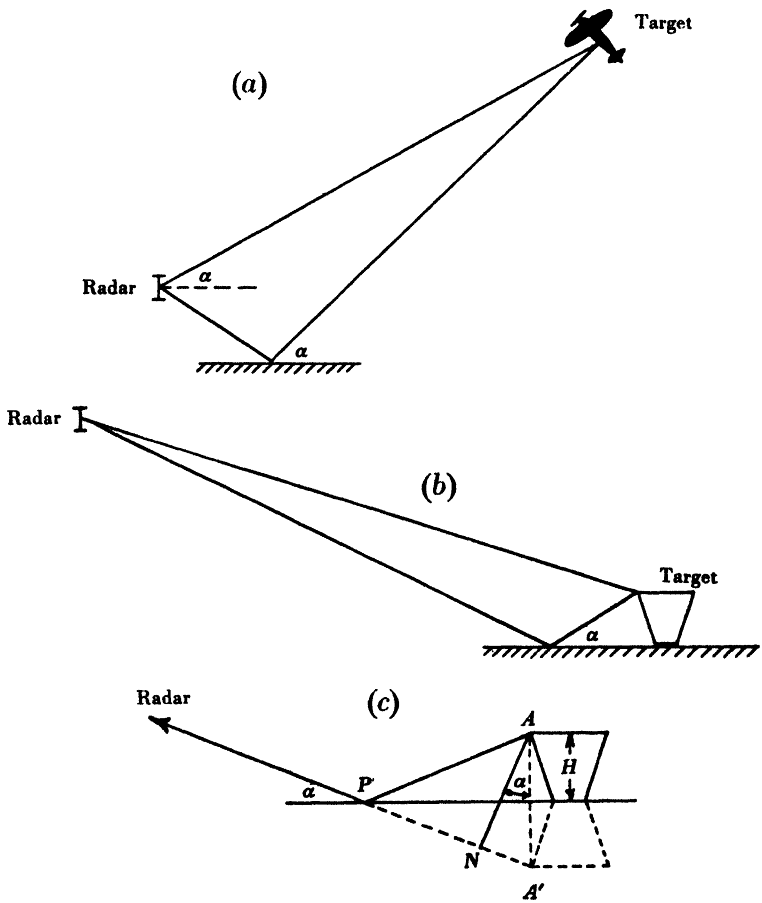


Fig. 3·9.

base to the top of the vessel. But if $2H \sin \alpha > \frac{1}{2}\lambda$, one or more interference maxima occur in the field strength falling on the target. A third possibility is that $\alpha < 0$, so that the radar is below the horizon at the target, and waves can only reach it by diffraction round the surface of the earth, although when this is the case the radar signals are generally (but not always) too small to be observed.

We shall therefore consider two or three zones, the 'near' zone when $4H \sin\alpha \gg \lambda$, the 'far' zone when $4H \sin\alpha \ll \lambda$, and the diffraction region, an extension of the 'far' zone when $\alpha < 0$.

In either of the first two zones, if the incident field strength neglecting the reflected wave is E_i, the field strength at height x above the sea will be (putting $k = 2\pi/\lambda$ as before)

$$2E_i \sin(kx \sin\alpha),$$

and if the incident energy-flux is F_i, the flux at height x will be

$$4F_i \sin^2(kx \sin\alpha).$$

If there is an element dS of echoing area at this height x, it will echo like an isotropic aerial emitting a power $4F_i \sin^2(kx \sin\alpha)\, dS$, but again there will be interference between a direct and a reflected ray. If the echoed radiation at angles of elevation α and $-\alpha$ is equal in magnitude and phase, we shall again obtain the same interference factor and the resultant flux at the radar receiving aerial will be

$$16F_i \sin^4(kx \sin\alpha)\, dS/4\pi r^2.$$

If we further assume that the different elements of echoing area radiate incoherently, so that we may sum the powers of the separate wavelets, we obtain a total flux at the radar receiver of

$$\frac{16F_i}{4\pi r^2}\int_0^H \sin^4(kx \sin\alpha)\, dS, \tag{3·8}$$

so that the effective echoing area is just sixteen times the integral in the above expression.

In practice it is found that the approximations we have made about non-coherence of phase of the wavelets from different parts of the target and the equality of $+\alpha$ and $-\alpha$ components are fairly well borne out, although unduly large signals are often obtained for broadside aspects, which may be due to some measure of coherence, and for bow aspects signals are rather small, due possibly to a tendency to scatter energy away from the incident direction. Otherwise it is found to be approximately correct to write $dS = 2b\, dx$, where b is the breadth of the vessel, which is substantially constant up to the height H of the top deck and very small above this. The factor 2 is empirical, but is what would be obtained if the scattering were isotropic into a hemisphere.

For the 'near' case, there are many lobes and gaps within the height H, and the average value of $\sin^4(kx\sin\alpha)$ is $\frac{3}{8}$, so that the effective echoing area is $\frac{3}{8}\times 16\int dS$, or $12bH$ to the approximation of the last paragraph, i.e. twelve times the projected area of the vessel. We write this quantity A_n, the effective echoing area for the 'near' case, so that $A_n = 12bH$. We see that in this case the effective echoing area is constant, so that the maximum range follows the same power law as in free space, which, as we shall see in §3·3, is that $P_s/P_m \propto r_m^4$.

For the 'far' case, on the other hand, $kx\sin\alpha$ is small, and we can approximate by writing $\sin^4(kx\sin\alpha)$ as $(kx\sin\alpha)^4$, thus obtaining an effective echoing area $A_f = 16(k\sin\alpha)^4\int_0^H x^4\,dS$. If the earth's curvature can be neglected, i.e. if $r/R \ll \alpha$, where R is the earth's radius, we may write $\sin\alpha = h/r$, where h is the height of the radar station above sea level, and we then obtain, writing $k = 2\pi/\lambda$,

$$A_f = \frac{256\pi^4 h^4}{\lambda^4 r^4}\int_0^H x^4\,dS,$$

and we see that the fundamental constant is no longer an area, but an integral of dimensions (length)^6. A concept of 'echoing volume' T, defined by

$$T^2 = \int_0^H x^4\,dS,$$

has been used to express this fact, but it must be noted that it is T^2 and not T itself which is proportional to the echoed energy flux density. We have in fact to replace the echoing area by

$$A_f = \frac{256\pi^4 h^4}{\lambda^4 r^4}T^2, \tag{3·9}$$

and it should be noted that this introduces an additional term r^4 into the expressions, so that the power law determining the maximum range becomes $P_s/P_m \propto r^8$. The 'far' zone is therefore sometimes referred to as the 'eighth power region'. It should be noted that on the approximation that $dS = 2b\,dx$, T^2 has the value $0{\cdot}4bH^5$, or $A_nH^4/30$, so that $A_f/A_n = (4\pi hH/\lambda r)^4/30$.

We see thus that for surface targets sufficiently near the radar there is a definable equivalent echoing area, and the signal strength

decreases with increasing distance (power varying inversely as the fourth power of the distance) according to the normal radar law. At greater distances is a transitional region, where the law can only be obtained by numerical integration of equation (3·8), before the 'eighth power zone' is reached. This zone ends as we approach the horizon as seen from the radar station, and we pass into the diffraction zone where the power/range law is steeper than the eighth power. In some cases there may be no clearly defined eight-power zone, if for values of α not much less than the critical $\sin^{-1}(\lambda/4H)$ the curvature correction to the relation $\sin\alpha = h/r$ is already appreciable.

Table 2 gives typical numerical values for A_n and T, being average values for oblique and stern aspects.

Table 2. *Average echoing constants—surface vessels*

Class of ship	Tonnage	H (m.)	A_n (sq.m.)	T (cu.m.)
Very small	20– 50	2·5– 4	50– 250	6– 40
Small	50– 200	4– 6	250– 750	40– 200
Trawler	200– 600	6–10	750– 4,000	200– 1,000
Coaster	600– 1,000	9–13	4,000–12,000	1,000– 3,000
Merchant vessel:				
Medium	1,000– 3,000	12–16	12,000–30,000	3,000–10,000
Large	3,000–10,000	16–20	30,000–50,000	10,000–20,000
Very large	Over 10,000	Over 20	Over 50,000	Over 20,000

N.B. Surfaced submarines usually fall into one of the first two classes, battleships into the last class.

3·3. Calculation of signal strengths and maximum ranges

We are now in a position to calculate signal strengths and maximum ranges of any given radar system, excepting only such cases as are affected by what are known as anomalous propagation conditions. These will be considered in the following section; meantime we should remark that in 'normal' propagation conditions there is a certain refraction of the rays towards the earth which is best represented by assuming for the radius of the earth R the value 5000 miles instead of the true 4000 miles (cf. § 3·4).

The calculation is made in several stages. First we calculate the incident flux density at the target by equation (3·4), then from the echoing area we deduce the power of the equivalent isotropic radiator, and applying equation (3·4) again obtain the reflected flux density, thus:

$$F_r = \frac{AF_i}{4\pi r^2}, \quad F_i = \frac{P_s G_s}{4\pi r^2},$$

whence
$$F_r = \frac{P_s G_s A}{16\pi^2 r^4}. \tag{3·10}$$

The 'available' power from the receiver aerial is then given by $A_a \times F_r$, or by equation (3·1)

$$P_r = \frac{G_r \lambda^2}{4\pi} F_r = \frac{P_s G_s G_r \lambda^2 A}{64\pi^3 r^4} \tag{3·11}$$*

This quantity can then be compared with the limiting noise sensitivity defined in § 2·5 and denoted by P_m; the maximum range is obtained by equating P_r and P_m, and is therefore

$$r_m = \sqrt[4]{\frac{P_s G_s G_r \lambda^2 A}{64\pi^3 P_m}}. \tag{3·12}$$

There are several points to note when using this formula. First of all, if there are any losses in the feeder systems, either in the sending or the receiving systems, these must be allowed for in defining G_s and G_r. Secondly, if reflected rays from the ground produce appreciable interference factors, these must also be allowed for in G_s and G_r. P_m must of course be correctly calculated for the pulse shape concerned and measured in terms of peak power if P_s is so measured, and if aerial noise is appreciable at the wavelength concerned this must be allowed for. Finally, should the target be a surface vessel, instead of allowing for the interference on reflexion in G_s and G_r, the considerations of the previous section must be used.

It will be seen that pulse length, pulse shape, etc., do not appear explicitly in equation (3·12). However, P_m is, other things being equal, proportional to the band width B, so that if we were to halve the band width and double the pulse length, a halving of the peak pulse power would appear to be allowable without diminishing the maximum range of the equipment. The reason why this cannot be

* Some authors (e.g. Scott, J. M. C., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 104, 1946) write the formula in the form

$$\frac{P_r}{P_s} = \frac{S_r S_s}{r^4} \frac{A}{4\pi\lambda^2} = \frac{S_r S_s}{r^4} \frac{L}{\lambda^2},$$

where S_r and S_s are the effective areas of the receiving and sending aerials respectively. The corresponding formula for one-way communication is

$$\frac{P_r}{P_s} = \frac{S_r S_s}{\lambda^2 r^2}.$$

done usually lies in considerations of range accuracy required, but it is important to realize that if this is already adequate, increasing the peak pulse power further may not be worth while if it involves shortening the pulse and widening the receiver band width.

3·3·1. *Aerial noise*

As mentioned in § 2·4·1, aerial noise is usually only important for wave-lengths longer than about 3 m. For most radar wave-lengths the ionosphere is substantially transparent, and we might expect the effective aerial temperature to approximate to that of interstellar space. In fact it is found that quite large amounts of radio noise in the 2-20 m. band reaches the earth from the plane of the Milky Way, the radiation being particularly strong in the direction of the centre of our galaxy (the direction of the constellation Sagittarius). The radiation is thought to arise from interstellar matter at a very high temperature, but the cloud of gas is not thick enough for black-body conditions to have been attained. Within any fairly narrow frequency band we may write (approximately) that the average noise temperature of the sky $T' = 14\lambda^3$ °K., i.e. 3000° K. at 6 m. wave-length and 24,000° K. at 12 m. The noise temperature of any aerial depends on its directivity—even with a non-directive aerial the earth's 'shadow' will introduce some diurnal variation in the noise (± 2 db. perhaps), while a highly directional aerial may have a noise temperature falling to say half the above figure but rising to four or five times this figure should it happen to point in a favourable direction.

We see that at $2\frac{3}{4}$ m. wave-length the mean noise temperature of the sky falls to 290° K., while for shorter wave-lengths the noise level presumably continues to fall, though measurements in this region are few. However, for the shorter wave-lengths the receiver noise factors are well above unity (cf. fig. 2·4) so that any noise from the aerial is swamped, and even if the aerial temperature fell to absolute zero very little improvement in performance would result. In using the formula $N-1+T'/T_0$ for the effective noise factor when the aerial temperature is T' (cf. § 2·4·1), it must be remembered that if the receiver aerial feeder introduces an attenuation of β, the formula becomes $N-1+\beta T'/T_0$. In the case (e.g. at 6 m. and above) where the third term in this expression is much larger than the others the β introduced here cancels a similar factor introduced

into G_r in equation (3·12), so that when aerial noise swamps set noise the range is unaffected by attenuation in the receiver feeder; attenuation in the sender feeder still reduces the range, however.

It must be noted that at wave-lengths longer than about 15 m. the above formula for noise level ceases to apply; the ionosphere begins to be opaque and other considerations arise, e.g. static from thunderstorms becomes important.

3·3·2. *Ground stations with site reflexions*

The case of ground stations with site reflexions is of some importance and merits further examination. If the sending and receiving aerials are identical, or have similar vertical polar diagrams and are sited at the same height, we can write a common elevation field-strength factor $\phi(\alpha)$ for the two aerials. For the case of a symmetrical array on a flat site ϕ is a product of two terms, viz. $f(\alpha)$, the directivity of the array itself, and a term of the form $2\sin(kh\sin\alpha)$, suitably modified if the reflexion coefficient is not -1, representing the effect of interference between the two rays. G_s and G_r are generally taken as the value for horizontal propagation (which is usually the direction of maximum gain for a symmetrical array), so that $f(0) = 1$. If, in addition, $\rho = -1$ and the first lobe maximum occurs at such a low angle that $f(\alpha)$ is still sensibly unity, the maximum range is given by $\phi = 2$. In other cases the maximum value of ϕ is less than 2. $\phi(0)$ of course always $= 0$.

To put this factor into equation (3·12), we need to remember that each gain, being a measure of power, is to be multiplied by the square of the field-strength factor, so that we must write $\phi(\alpha)^2 G_s$ in place of G_s and $\phi(\alpha)^2 G_r$ in place of G_r. We obtain the result, therefore,

$$(r_m)_{\text{int.}} = (r_m)_{\text{f.s.}} \times \phi(\alpha), \tag{3·13}$$

where the subscript f.s. refers to the free space condition and int. to the case with interference from the reflected ray. We see, for example, that in the best case, the interference range is just twice the free-space range calculated from the unmodified equation (3·12), the conditions for this to hold being those stated at the end of the previous paragraph.

We may extend this result to unlike aerials if we write $\phi(\alpha)$ for the geometric mean of the elevation field-strength factors of the

two arrays. Unless both arrays happen to give the value of 2 to their separate ϕ-factors for the same value of α, the geometric mean ϕ will have a maximum value which is less than 2. Otherwise the analysis is the same as before.

In computations the value of α must be obtained allowing for the curvature of the earth (using $R = 5000$ miles, see above and § 3·4). The formula usually used is $\sin\alpha = h/r - \frac{1}{2}r/R$, or if h is in feet, r in miles, and R is taken as 5280 miles, $5280 \sin\alpha = h/r - \frac{1}{2}r$. Alternatively, a graphical plot on curvilinear coordinates may be used, as shown in fig. 3·10, which is an example of a coverage diagram

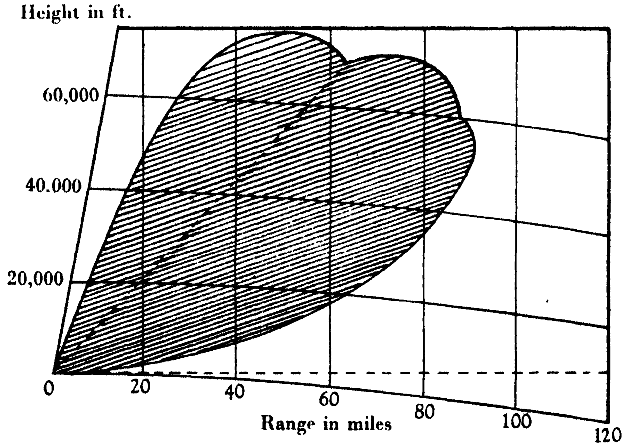


Fig. 3·10. Example of coverage diagram using curvilinear coordinates.

drawn for the vertical plane, a type of diagram which is often useful for indicating the possibilities and limitations of a ground radar station. The curve drawn is a locus of constant field strength, usually chosen to represent a just-detectable signal from a target of fixed echoing area (e.g. a medium-sized aircraft).

Below the maximum of the first lobe this theory is only approximate, being affected by diffraction of the wave around the curved surface of the earth. The range on targets well below the first lobe maximum is usually increased somewhat, but the effect is difficult to calculate and is small under normal conditions for most stations, particularly for short-range stations. This is just the region where anomalous propagation (vide next section) is most felt. It should be

noted, however, that in so far as our approximation is valid—e.g. for a flat earth—we have for small α, $\phi \propto \alpha \propto h/r$, and putting this into equation (3·11) we obtain

$$P_r \propto \frac{P_s G_s G_r \lambda^2 A h^4}{64\pi^3 r^8},$$

so that we see that for aircraft flying well below the maximum of the first lobe we obtain approximately an eighth-power law connecting received power and range, as we did in the 'far' zone for surface targets.

3·3·3. *Performance index*

From equations (3·11) and (3·12) we see that the ratio of actual signal power (P_r) to the minimum detectable (P_m) is given by $(r_m/r)^4 = S^2$ if S is defined as the ratio of the signal amplitude to the minimum detectable. Thus for any echo the quantity

$$\frac{S^2 r^4}{A} = \frac{P_s G_s G_r \lambda^2}{64\pi^3 P_m} = I_p, \quad \text{(say)},$$

should be constant, depending on the equipment only, except for the site reflexion cases considered in the last section, where it will also depend on the angle of elevation, and excepting also surface-vessel detection beyond the 'near' zone. This quantity I_p (or its logarithm, usually expressed in decibels) we may call the performance index of an equipment—for the site reflexion case the index is usually defined for the angle of elevation giving the maximum value. In our units I_p is measured in (metres)^2, but other units are also used; a common form in decibels is

$$I_{\text{db.}} = 10 \log (P_s G_s G_r \lambda^2 / \Pi N B) + 68,$$

with P in kW., λ in cm., and B in Mc./sec.; some writers introduce an alternative form of the performance index in which the factor λ^2 is omitted.

The concept of a performance index is of some importance, but the diversity of units used is such that the basis adopted should be clearly explained whenever the concept is used. The authors' preference for a consistent system of units such as the metre-kilogram-second system is unfortunately not generally shared.

3·3·4. *Numerical examples*

(1) Predict the maximum range on a medium aircraft (say $A = 13$ sq.m.) of a 1·5 m. equipment with a sender power of 100 kW.,

noise factor 3·4, band width 4 Mc./sec. and perception factor 1·5 using for both sending and receiving a broadside array of 32 dipoles with a reflector and a 1 db. loss in the feeders for both sending and receiving. The equipment is on a cliff site ($\rho = -1$) and the aerial has substantially maximum gain at the elevation of the first lobe maximum.

Here $P_m = 8{\cdot}1 \times 10^{-14}$ W. (cf. § 2·6, example 2), $P_s = 10^5$, $G' = 64$ less 1 db. for both aerials, whence

$$G_s = G_r = 1{\cdot}64 \times 64 \times 0{\cdot}794 = 83{\cdot}5.$$

Therefore from equation (3·12)

$$(r_m)_{\text{f.s.}} = \sqrt[4]{\frac{10^5 \times (83{\cdot}5)^2 \times (1{\cdot}5)^2 \times 13}{64 \times \pi^3 \times 8{\cdot}1 \times 10^{-14}}}$$

$$= 1{\cdot}06 \times 10^5 \text{ m.} \quad \text{or} \quad 106 \text{ km.}$$

Now under the conditions stated ϕ has a maximum value of 2, whence $(r_m)_{\text{int.}} = 2 \times 106$ km. or 132 miles.

(1 *a*) Compare with this the predicted range for the same radiating system and receiver but used for one-way transmission instead of radar. Even if the receiving aerial were only a $\frac{1}{2}\lambda$ aerial, we have

$$F = \frac{P_s G_s}{4\pi r^2} \quad \text{and} \quad P_m = A_a F,$$

where $A_a = 1{\cdot}64\lambda^2/4\pi$, or $\lambda^2/8$ (eqn. (3·1 *a*)), whence

$$8{\cdot}1 \times 10^{-14} = \frac{10^5 \times 83{\cdot}5 \times 1{\cdot}5^2}{32\pi (r_m)^2_{\text{f.s.}}},$$

giving $(r_m)_{\text{f.s.}} = 1{\cdot}5 \times 10^9$ m. or about one million miles. This figure serves to emphasize that a different order of magnitude of sender power and receiver sensitivity is necessary for radar than suffices for most ordinary communications equipment.

(2) An airborne 10 cm. radar has a sender power of 50 kW. and a limiting noise sensitivity of 10^{-12} W. (cf. first example of § 2·6). The aerial is an 18 in. diameter paraboloid, the effective aperture being three-quarters of the actual aperture, and is used for sending and receiving, using beam-switching (cf. Chapter 5) so that the gain at the working point is 60 % of that in the lobe maximum. What maximum range would be expected on a medium bomber (for which $A = 20$ m.^{2})?

Here $P_s = 5 \times 10^4$, $P_m = 10^{-12}$,

$$G_s = G_r = 0{\cdot}6 \times \tfrac{3}{4} \times \frac{4\pi\,(\text{aperture})}{\lambda^2} = 1{\cdot}8\,\frac{\pi^2(9 \times 2{\cdot}54)^2}{10^2},$$

so that using equation (3·12) and simplifying

$$r_m = 2{\cdot}54 \times 9 \times 10^3 \times \sqrt[4]{(1{\cdot}8^2 \times \pi/64)}$$
$$= 1{\cdot}45 \times 10^4\,\text{m.} \quad \text{or} \quad 14{\cdot}5\,\text{km.} \quad \text{or} \quad \text{about 9 miles.}$$

(3) A C.H. station has a sender power 200 kW. at 12 m., feeding an array having a gain of 12 times that of a $\frac{1}{2}\lambda$ aerial. The receiving aerial has a gain of twice a $\frac{1}{2}\lambda$ aerial, and the centres of both are 48 m. above a perfectly flat reflecting site. An aircraft of $A = 10\,\text{m.}^2$ at 60 miles range and 15,000 ft. high gives a signal of 2·5 times the amplitude of the minimum detectable signal, the band width being $\frac{1}{2}$ Mc./sec. and the perception factor 1·5. Calculate the equivalent noise temperature of the aerial, assuming receiver noise negligible. First we find α, using

$$5280 \sin\alpha = h/r - \tfrac{1}{2}r,$$

i.e.

$$= 15{,}000/60 - 30 = 220,$$

whence $\sin\alpha = 1/24$. Also $k = 2\pi/12 = \frac{1}{6}\pi$. Therefore

$$\phi = 2\sin(kh\sin\alpha) = 2\sin(48/24 \times \tfrac{1}{6}\pi = \tfrac{1}{3}\pi) = \sqrt{3},$$

also

$$60 \text{ miles} = 9{\cdot}64 \times 10^4\,\text{m.}, \quad G_s = 1{\cdot}64 \times 12 = 19{\cdot}7, \quad G_r = 3{\cdot}28,$$

whence

$$P_r = \frac{P_s G_s G_r \phi^4 \lambda^2 A}{64\pi^3 r^4} = \frac{2 \times 10^5 \times 19{\cdot}7 \times 3{\cdot}28 \times 9 \times 12^2 \times 10}{64\pi^3 \times (9{\cdot}64)^4 \times 10^{16}}$$
$$= 9{\cdot}8 \times 10^{-13}\,\text{W.};$$

also $P_r = S^2 \Pi k T' B = 2{\cdot}5^2 \times 1{\cdot}5 \times 1{\cdot}37 \times 10^{-23} \times \frac{1}{2} \times 10^6 \times T'$

$$= 6{\cdot}17 \times 10^{-17} T'\,\text{W.},$$

whence $T' = 9{\cdot}8/6{\cdot}17 \times 10^4 = 16{,}000^\circ\,\text{K.}$

N.B. Notice that to equal this noise level a receiver noise factor of $16{,}000/290 = 55$ would be necessary. Any ordinary good receiver will therefore contribute negligible noise.

(4) (*a*) What is the field strength due to the C.H. sender at this aircraft, and (*b*) what power radiated from a non-directional aerial

in this aircraft, 100% noise-modulated over a band width of 2 Mc./sec., would be necessary to increase the received noise level so that 40 miles became the maximum range of the station on an aircraft having $A = 10$ m.^{2}?

(a) $F = E^2/120\pi = P_s G_s \phi^2/4\pi r^2$, therefore

$$E^2 = 30 P_s G_s \phi^2/r^2 = 30 \times 2 \times 10^5 \times 19{\cdot}7 \times 3 \div (9{\cdot}64 \times 10^4)^2,$$

$$E = \sqrt{(18{\cdot}5 \times 19{\cdot}7 \times 10^6)}/9{\cdot}64 \times 10^4 = 1{\cdot}88/9{\cdot}64 = 0{\cdot}195 \text{ V./m.}$$

Notice that although we normally work in terms of flux densities, the field strength is readily obtained when required.

(b) The maximum range of the station at present is $60 \times \sqrt{2{\cdot}5}$ miles at elevation $\sin^{-1} 1/24$, and $2/\sqrt{3}$ of this in the lobe maximum ($\phi = 2$); i.e. $120\sqrt{(2{\cdot}5/3)} = 109{\cdot}3$ miles. If this is reduced to 40 miles, the noise power must be increased as r^{-4}, i.e. as $3^4(2{\cdot}5/3)^2 = 56{\cdot}25$ times. A jamming noise of $55{\cdot}25 \times 9{\cdot}8 \times 10^{-13}$ W. is therefore required. Let the jammer power $= P_j$, then $0{\cdot}25 P_j$ will lie within the receiver pass band. The receiving aerial has $G' = 2$ so that its $A_e = \lambda^2/4$ (cf. eqn. (3·1 a)) $= 36$ m.^{2}, so that

$$\phi^2 \times 36 \times 0{\cdot}25 P_j/4\pi(9{\cdot}64 \times 10^4)^2 = 55{\cdot}25 \times 9{\cdot}8 \times 10^{-13},$$

and putting $\phi^2 = 3$ as above, we obtain $P_j = 0{\cdot}23$ W.

3·4. Propagation conditions in the earth's atmosphere

The refractive index of the atmosphere depends on its density and humidity; the water molecule, being highly polar, makes a relatively large contribution for its density. Approximately, for any frequency within the radar band, the refractive index μ is given by the expression*

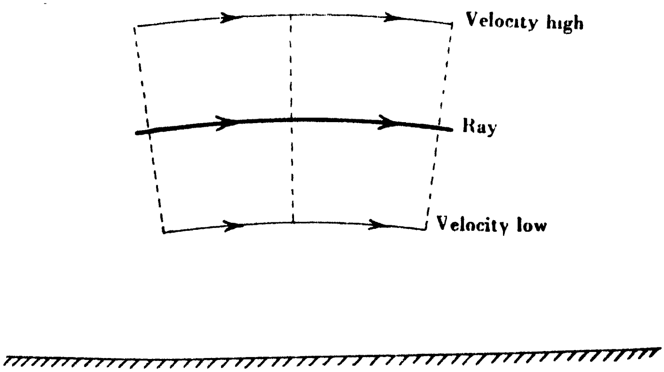
$$\mu - 1 = \frac{80}{T} \times 10^{-6}\left(p + \frac{4800}{T} w\right),$$

where p is the atmospheric pressure in millibars, w the partial pressure of water vapour in millibars, and T the absolute temperature (° K.).

The absolute value of this index is only required when very accurate ranging is necessary; for normal purposes it is the bending of the rays due to gradients of refractive index which is more important. Under average or normal conditions this index decreases

* England, Crawford and Mumford, *Bell Syst. Tech. J.*, vol. 14, p. 369 (1935).

as we go upwards in the atmosphere; the velocity of radio signals therefore increases with height, causing the rays to bend downwards towards the earth (fig. 3·11). The ray curvature in average conditions is some one-fifth of the curvature of the earth; if therefore we distort our coordinate system until these rays are straight the earth's curvature is reduced to some four-fifths of its true value, corresponding to an effective radius of 1·25R or about 5000 miles. As mentioned in § 3·3·1, this figure is often used in radar calculations of angle of elevation of distant targets.



----- Successive wave fronts

Fig. 3·11. Illustrating bending of ray towards the earth.

This convention can, however, only approximately represent actual conditions, which may vary a great deal from the average according to the state of the weather. It is usually referred to as 'orthodox' propagation, but in some meteorological conditions a strikingly different state of affairs arises, known as 'unorthodox' or 'anomalous' propagation. This occurs when the ray curvature near the earth's surface becomes greater than the radius of the earth itself. The meteorological conditions under which this occurs are fairly well defined, being usually associated with a temperature inversion, but for details original papers* should be consulted. In England the phenomenon occurs moderately frequently, being most likely on fine summer evenings.

Unorthodox propagation is characterized by the formation of a 'duct' close to the earth's surface in which the waves are trapped as

* E.g., Booker, H. G., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 69 (1946).

if in a wave guide. The upper surface of the wave guide is not, as it generally is, a conductor reflecting the waves; instead the refraction bends the waves concerned back into the duct (fig. 3·12). Just as in a conventional wave guide there is a critical wave-length above which propagation along the guide does not occur, so the duct can only trap energy if the wave-length is short enough—longer waves behave relatively normally. The duct width is defined as the height from the ground to the ray whose curvature is just equal to that of the earth (cf. fig. 3·12). The critical wave-length is related to the duct width as shown in table 3; the relation is not the simple one

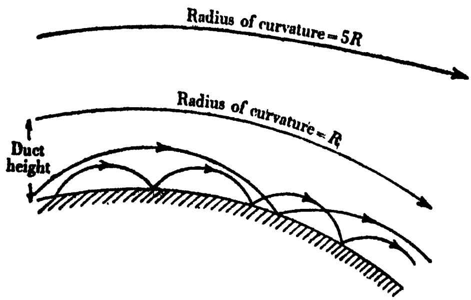


Fig. 3·12. Illustrating duct formation.

obtaining in the metallic guide, since the upper surface can only bend the rays back gradually, so that the critical wave-length is much smaller than for a metallic guide of the same width.

Table 3

Wave-length (m.)	Min. duct height (m.)
1·5	230
0·5	110
0·1	36
0·03	17
0·01	8

Very narrow ducts are in fact found to exist quite frequently, so that for wave-lengths of the order of 1 cm. 'unorthodox' conditions might be the rule rather than the exception. However, the wider the duct the less frequently it is formed, so that in England 'unorthodox' propagation on 5–10 m. may occur less often than once a year, while on 50 cm. it might occur perhaps on 10 days of the average year for several hours of each day.

The result of the formation of such a duct is that the field strength close to the earth's surface attenuates much more slowly than under orthodox conditions, and the horizon ceases to provide any limitation at all. The effects are most marked if the radar aerial is sited low enough to lie within the duct. The field strength above the duct is somewhat reduced, but the duct only traps radiation leaving the aerial at a suitable angle, so that quite good field strengths are still obtained outside the duct, and high-angle targets are still detectable. However, the effect increases enormously the visibility to radar of low objects at considerable distances. Thus the coast of Holland, which is far under the horizon from stations on the east coast of England, may be detected clearly under unorthodox propagational conditions. Sometimes objects are detected at such a great range that their responses follow the next pulse of the radar sender, and may be erroneously interpreted as echoes from that pulse at much shorter ranges. In the tropics cases have been authenticated of radar echoes being received from ranges well above 1000 miles due to duct conditions.

3·5. Choice of constants for a given radar equipment

By way of summarizing the considerations of this and the previous chapters we propose to survey briefly the factors governing the choice of frequency and other constants for a given radar purpose. These factors naturally depend on the state of development of the art—so that in 1938, for example, only frequencies up to about 100 Mc./sec. were in the picture. There was therefore at this date substantially no alternative to erecting large towers to carry the aerials of coastal aircraft-detecting stations.

At the present time we can attain pulse powers P_s of the order of 200 kW. over most of the frequency band from 30 to 10,000 Mc./sec., and 1 MW. at the longest wave-lengths and at 3000 Mc./sec. and thereabouts, while the band from 1000 to 3000 Mc./sec. is relatively undeveloped. P_m, on the other hand, tends to rise somewhat with increasing frequency (cf. § 2·6). Considering equation (3·12), therefore, we see that the governing factors, apart from the factor P_m, are the terms $G_s G_r \lambda^2$. If, as in an airborne installation, the aerial aperture is strictly limited, this fixes the product $G_s \lambda^2$ (cf. eqn. (3·1)), but reducing the wave-length allows G_r to increase as λ^{-2}, so that the range tends to increase as $\lambda^{-\frac{1}{2}}$. On the other hand,

in circumstances where the gains are fixed, e.g. if beams of a given angular width are required, the wave-length factor goes the other way and a long wave-length is indicated, providing the aerial size required is not prohibitive. There is thus usually a rough optimum wave-length for any given function.

In cases where site reflexions are important, or for the detection of surface vessels in the 'far' zone, there are additional factors to be taken into account, and sometimes operational considerations are also involved. Thus for a ground coastal reporting station 100–200 Mc./sec. might be a satisfactory frequency if the disadvantage of having alternate lobes and gaps due to sea reflexions was not important, although the use of a shorter wave-length might enable an aerial to be used which was more directive in the vertical plane so as to give a satisfactory coverage diagram without gaps. For ground stations, too, considerations of propagational conditions to be expected at the site may determine the most suitable wave band—for example, in some tropical locations 'duct' conditions would be almost permanent for centimetre wave-lengths, and this might in some circumstances make their use undesirable.

The restriction that the beam shall not be too narrow, as it might if too short a wave-length were used, is sometimes of importance. With a very narrow beam difficulty may be encountered in initially locating a particular target. In other types of system, where the beam is required to scan in two angular dimensions, another limitation arises; in order to obtain a good response from any target with a reasonably low perception factor, it is essential that a certain number of consecutive pulses should be obtained from it during the time when the beam sweeps over it. These considerations are treated in more detail in Chapter 7.

How all these factors, and in particular the choice of wave-length, work out in practice is probably best realized from a consideration of the various radar equipments listed in Chapter 9. In general, it will be seen that wave-lengths in the centimetre band have found favour where the aerial size is severely restricted, as in airborne equipments, while for ground stations the longer range of the metric wave equipments competes with the higher accuracy of the centimetre ones, and the choice is a more open one.

Chapter 4

RADAR RANGING

We have already indicated that distance determination using radar is essentially a process involving the measurement of the time delay of the returned pulses. It is proposed in this chapter to discuss how this measurement is made and give details of the order of accuracy attainable.

The time delay is equal to $2r/c$, where r is the range of the radar target and c is the velocity of propagation of electromagnetic waves, so the first question is how accurately we know this latter factor. The experimental determination of the velocity of propagation of electromagnetic waves has been largely due to Michelson and his collaborators, who, using optical methods, have obtained a value for c of $299{,}774 \pm 11$ km./sec. The value of c may therefore be taken as known to an accuracy of about 1 part in 3×10^4. According to the classical theory of electromagnetism, the velocity in free space should be independent of frequency; further, according to the special theory of relativity any measurement of this velocity should be independent of the velocity of the observer. It can therefore be concluded that the velocity of propagation of radar waves in free space will be the same as that of light in free space.

In our case, however, we are not concerned with the velocity of propagation in free space so much as with the velocity of propagation in the earth's atmosphere. This velocity is less than the free-space velocity by a factor μ, the refractive index of the earth's atmosphere for radio waves. In terms of the temperature of the air in degrees Kelvin (T), the total pressure of the atmosphere in millibars (p), and the partial pressure of water vapour in the atmosphere also in millibars (w), the refractive index is given by*

$$\mu - 1 = \frac{80}{T} \times 10^{-6}\left(p + \frac{4800w}{T}\right). \tag{4·1}$$

There are, of course, day-to-day variations in the values of p, T and w, but for average conditions we can calculate the variation

* Cf. § 3·4.

of velocity of propagation with height. The results are given in table 4. It will be seen that the mean velocity of propagation of radio waves at ground level differs from that in free space by about 1 part in 3000. Estimates have been made of the daily variation of p, T and w, and using this data in equation (4·1), it is found that the maximum variation is of the order of 1 part in 10^4 at low altitudes and decreases to 1 part in 4×10^4 above 5000 ft.

Table 4

Height in feet	0	10,000	20,000	30,000
Mean velocity in km./sec.	299,680	299,710	299,725	299,740

The velocity of propagation of radio waves in the earth's atmosphere has also been determined directly. Perhaps the most accurate measurements are due to Smith, Franklin and Whiting.* They used what was effectively a radar method and timed the transmission of radio pulses between two ground stations. Various frequencies between 20 and 60 Mc./sec. were used, and they obtained an average value of $299{,}705 \pm 50$ km./sec. They conclude as a result of their experiments that this result is accurate to 1 part in 6000. It is thus apparent that the accuracy of radar ranging is limited by the uncertainty of the precise value of c to use for the particular propagation path involved, but that since c is usually known to 1 part in 6000, the range measurement can be made to this order of accuracy if instrumental errors are negligible.

4·1. Method of measurement of range

The problem of range determination can now be summarized as one of measuring a time interval, possibly up to 2000 μsec., occurring between the leading edge of the sender pulse and the leading edge of the received pulse. A radar ranging system will therefore require (i) some means of displaying the sender pulse and the reflected pulse on a suitable time base, and (ii) some standard of time (or frequency) to which reference may be made. Furthermore, the display system must be synchronized to the sender. To do this it is usual to generate in a low-impedance source a sharp-edged synchronizing pulse, one of whose edges is the reference instant of the system. In some cases

* U.R.S.I. Convention Paper, Paris, October 1946.

this is generated separately and drives both the display timer* and the modulator. In other cases the sender delivers a trigger pulse to the display timer at the beginning of each pulse of radio-frequency energy, and all the other operations of the system are timed from this trigger. Systems of the first type are usually called *Master-oscillator Systems*, since the master oscillator in the timer determines the pulse-repetition frequency, and from its output further wave forms are produced for controlling the entire system. Systems of the second type are called *Self-synchronous Systems*. Whichever method is used, however, the time of rise of the synchronizing pulse determines the maximum precision of timing that can be achieved.

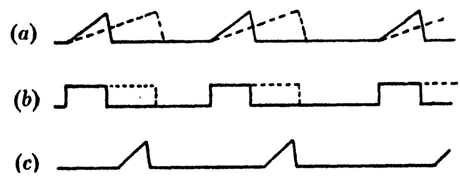


Fig. 4·1. Time-base wave forms.

It is also necessary to generate a suitable wave form for the display timer. In the simplest case this will be a linear time base, possibly of adjustable length, deflecting a cathode-ray tube spot, and with this will be associated a black-out wave form to illuminate the tube only during the forward stroke of the time base. These wave forms are shown as (*a*) and (*b*) in fig. 4·1. The echo signals are usually displayed by deflexion at right angles to the time-base length, and in some cases alternative time-base lengths are provided to allow viewing of targets at long or short range. This type of display is commonly used for search radars, although the accuracy attainable is very definitely limited. For example, with a 10 cm. trace, it is obviously very difficult to estimate to better than $\pm 0{\cdot}5$ mm. on the tube face, and for a range display of 0–30,000 yards, this limits the accuracy to something of the order of ± 150 yards. Some systems obtain an improvement by using circular or spiral time bases with signal deflexions radially. For precision gunnery, however, the highest range accuracy is desirable. One method is to extend the time base to five or six times its normal length and to display only

* Cf. fig. 1·4.

a small portion on the tube face. A delayed fast time base such as is shown in fig. 4·1 (*c*) may also be used. This is often obtained by amplifying a portion of the long time base, but in some cases it is more convenient to trigger a fast time base by means of a delayed pulse.* A 'bias' potentiometer may be used to bring any selected part of the trace to the centre of the cathode-ray tube screen. This potentiometer is accurately made and calibrated, and since bringing the selected echo to a cross-wire on the centre of the tube requires a definite bias voltage, the position of the potentiometer for this setting gives a measure of the range of the target. This method is often referred to as the 'cross-wire' method. The range data may be

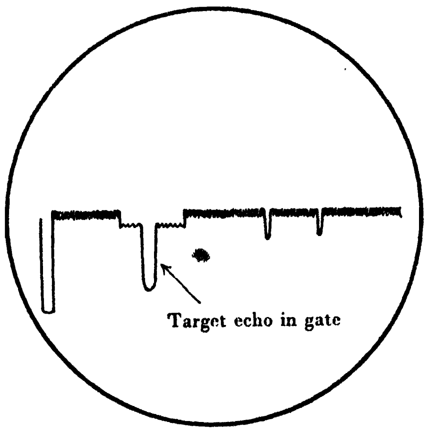


Fig. 4·2. Display of ranging system with expansion of trace.

transmitted to the computor (cf. §4·1·1) either as a voltage (proportional to the range) or mechanically as a movement along a scale (also proportional to the range).

In some cases it is convenient to display the whole of the long time base plus a small speeded-up portion. This is shown in fig. 4·2. The technique is then as before, the position of the speeded-up portion (or gate) is adjusted so that the selected echo is in the

* A delayed pulse of this kind is often employed to mark a particular echo using brightness modulation, by adjusting its delay to equal that of the echo. This is of value in correlating the range data of a particular echo, with the direction-finding data for that echo. Such a pulse is called a 'strobe' or a 'gate'.

gate and its leading edge corresponds to the beginning of the speeded-up portion. The position of the gate potentiometer, which controls the delayed pulse, then provides a measurement of the range of the target.

In many equipments, and this applies particularly to search radars, the theoretically possible range accuracy is not in fact required. This is fortunate because search radars which are generally concerned with the determination of azimuth as well as range usually employ a range azimuth or P.P.I. type of display, and the use of this in place of a range-amplitude display usually involves a definite sacrifice in the range-reading accuracy.

4·1·1. *Range measurement methods for use with computors*

For some purposes a form of automatic transmission of the range (and sometimes the range rate) from the radar display to a computing mechanism is necessary. If, for instance, the range and elevation of the target are determined separately, it is useful to be able to feed the range data automatically to a height computor. Such a system is shown in fig. 4·3. To feed the range information to a computor or predictor the operator is required to adjust the range dial continuously so that the cursor across the face of the cathode-ray tube coincides with the leading edge of the echo. In this case since the range index is mechanical, it is necessary to ensure that the calibration law is correct, and this, of course, involves periodic checks with the calibrator signals.

In the case of fire-control radar, continuous range and range-rate information is often required to allow the calculation of the 'future position' of the target. In practice, it is a great advantage if some form of rate-aided transmission unit can be employed for transmitting the rotations of the range dial to the predictor, since by this means more accurate 'following' of the target can be achieved.

In some cases, a completely automatic radar is required, i.e. it is required to feed range, bearing and elevation data into the predictor without the intervention of any operator. This means that a range marker or strobe is required which is maintained in alinement with any selected echo automatically. Similarly, the axis of the aerial is required to maintain a direction along the line of sight automatically. These two processes are known as 'auto-ranging' and 'auto-

aiming' respectively. Auto-ranging and auto-aiming are of special importance when the target is accelerating, and it is in the case of close-range gunnery actions that the accelerations are greatest and where this technique is of greatest value. The basic principle of auto-ranging is that of using a range marker or walking strobe* which is made to traverse the cathode-ray tube trace until it embraces the selected echo, and then to follow its movements. This is brought about by generating a misalinement voltage which is applied to the 'strobing' circuit in such a sense as to tend to reduce the misalinement. Auto-follow systems of this type are very complex and require

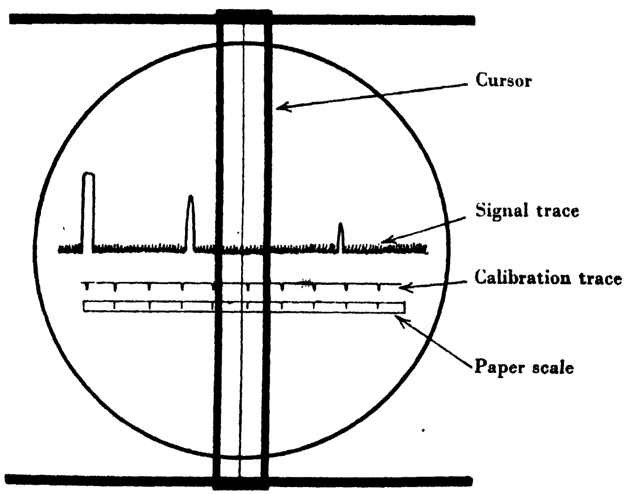


Fig. 4·3. Mechanical range index.

very careful design to prevent instability. Nevertheless, they have found many applications, viz. army gun-laying equipments, naval precision fire-control sets, and air interception equipments. In fact, any equipment which needs to track one selected target only usually benefits considerably by being fitted with auto-following. This is partly because the operator is dispensed with, but more particularly because auto-following eliminates all the errors which are due to a variable time constant and inconsistency on the part of the operator.

* Williams, F. C. and Ritson, F. J. U., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 318 (1946).

4·2. Calibration

4·2·1. *Frequency as a standard*

The precision of the time-base circuits outlined in the previous section is limited by component variations. It is therefore necessary to have an accurate standard for calibration purposes. The stable L.C. oscillator, or better still the crystal-controlled oscillator, makes a suitable standard. An oscillator of 93·117 kc./sec. is frequently used, and provides one cycle for each mile of effective displacement of the cathode-ray tube spot. The sinusoidal wave form is unsuitable for calibrating purposes, and it is converted into range markers as

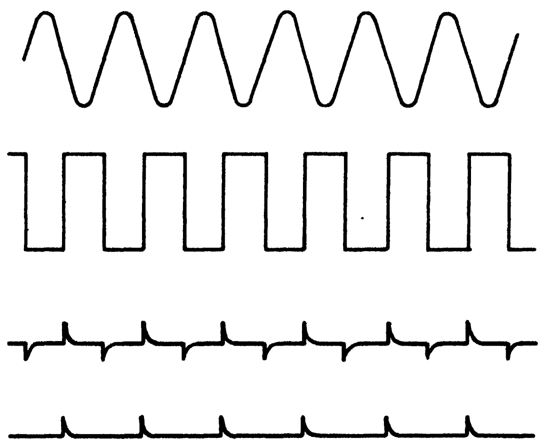


Fig. 4·4. Calibration wave forms.

shown in fig. 4·4. The sine wave is first converted into a steep-sided square wave, then into a series of positive and negative pips, and finally the positive pips are sharpened and the negative pips removed. For details of the circuit techniques employed to produce these changes we must refer the reader to the companion volume in this series, *Radar Circuit Techniques*, by Dr F. C. Williams.

It is often convenient to generate a calibrator trace immediately below the signal trace in the quiescent period between the end of the range trace and the next sender pulse. The calibrator signals are then presented continuously. When the interval between successive traces is too small to allow a complete calibration trace, one may arrange to substitute a calibration trace for the main display when

required at the turn of a switch. For example, we may employ a continuously running L.C. or crystal-controlled oscillator for calibration, and control the timing of the sender pulses from the same oscillator at a subharmonic frequency. However, the use of a system of this kind has certain operational disadvantages in that the pulses repeat themselves at exactly equal intervals of time, and the system is then more susceptible to jamming. One of the most important methods of countering interference, whether it be deliberate jamming by the enemy, or interference incidental to the operation of a large number of equipments on neighbouring (or the same) frequency channels, is to produce a 'jitter' or irregularity of the pulse-recurrence frequency. For this reason, although sinusoidal oscillators have been used for generating the synchronizing wave forms (cf. §4·1), some form of multivibrator whose frequency is either incidentally or deliberately made unstable is more often used. Calibration is then possible if each synchronizing pulse also triggers a damped oscillation in the calibrating L.C. circuit, but a better arrangement is to initiate the time base, during calibration, from special pulses generated at a subharmonic of the continuously running calibration oscillator. The method is, however, only applicable if the circuits are such that the time-base speed is unaltered by the change of synchronizing pulses.

Other standards which have found application as calibrators are (i) valve-controlled tuning forks and (ii) magnetostriction oscillators. Equipment of the former kind was used both by the Germans and the Italians but did not give as great an accuracy as the other methods.

4·2·2. *Supersonic velocity as a standard*

An alternative and perhaps more direct method of measuring the range depends on the use of a supersonic cell. The principle is that of using two pulses, a supersonic pulse and a radio pulse; the pulses are synchronized in time, and the received supersonic pulse after traversing a column of liquid is matched in time with the radio pulse which traverses a path from the sender to the target and back to the receiver. The two pulses are displayed on the same cathode-ray tube and the length of the liquid column is adjusted until the two pulses are received back at the same instant of time. A high-speed time base allows this to be done with great accuracy. When the

adjustment has been made the range of the radar target is equal to half the distance traversed by the supersonic wave multiplied by the scaling factor *k*. *k* is, of course, equal to the ratio of the two velocities of propagation.

The velocity of propagation of supersonic waves in water is about 1500 m./sec., and therefore the cell can be of quite moderate dimensions (e.g. a target at a range of 20 miles requires a water path of about 30 cm.). A vertical cylindrical water cell is employed in most cases. It uses a quartz crystal at its lower end and a plunger having a plane end which acts as a reflector. The quartz crystal is used for both sending and receiving and since the water is traversed twice only half the length is required.

The main advantages of the supersonic range scale are that it is continuously variable and self-calibrating. In addition, since the method involves the optical coincidence of two pulses on the display and the measurement of a length (in the cell) which may be carried out with a micrometer adjustment, high accuracy can be expected. One disadvantage is the fact that the velocity of supersonic waves varies markedly with temperature, and this seriously limits the accuracy unless thermostatic control is used.

4·2·3. *Electronic range marker*

We have already noted that where high accuracy is of secondary importance, it is usual to read the range of the target directly from the cathode-ray tube. In other cases a range knob is employed, which is linked with a mechanical cursor whose position is adjusted to correspond with the leading edge of the target echo. A magnified display allows this setting to be made with high accuracy, although since reference must be made to the calibration scale and, in general, interpolation between the calibration pips is necessary, the accuracy of the range measurement is obviously limited.

A more satisfactory method which is employed in the case of certain long-range gunnery sets and also in the case of a number of the radar navigational systems, uses a range index which is produced electronically. In this way the difficulty and uncertainty of matching the electronic range scale to the mechanical one is avoided. Fig. 4·5 illustrates this system. Two traces are used, on one of which the echo signals are displayed and the other is a calibration trace. A small portion of the signal trace (corresponding to the distance between

two calibration pips, say to 1000 yards) is speeded up, and since the calibration trace is developed by the same circuits, this also has the speeded-up portion corresponding to the same limits. Up to this point, therefore, the accuracy of the range determination depends on the scale of the picture as in previous methods. However, the range hand wheel which controls the position of the expanded portion of the trace is geared to a continuously adjustable phase shifter in the output of the crystal-controlled oscillator from which the calibration pips are derived. One complete rotation of the phase shifter corresponds to a change of phase of 2π or one complete

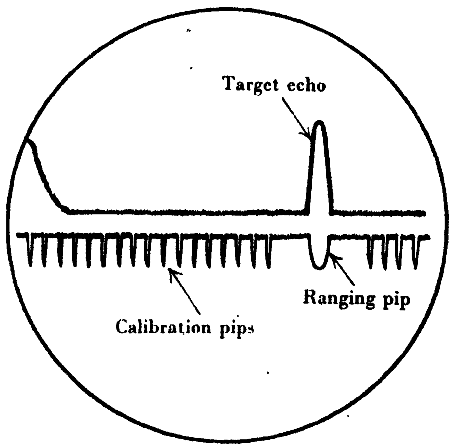


Fig. 4·5. Precision ranging system.

period of the crystal-controlled oscillator. Therefore, if one of the calibration pips is set opposite the ground wave and the hand wheel is rotated until this pip is set opposite the selected echo, the number of rotations of the hand wheel gives an accurate measure of range of the target. Thus, if the crystal-controlled oscillator produces 1000 yard pips, and it is necessary to rotate the hand wheel through 12 complete revolutions, and 0·65 of a complete rotation, the range of the target is evidently 12,650 yards.

Methods of this type are capable of very high accuracy, and, for one particular naval version of this equipment, an instrumental accuracy of about ± 5 yards is claimed.*

* Laws, C. W., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 423 (1946).

4·3. Radar ranging with range-elevation correlation

In certain cases, e.g. in the case of the radar range attachment to the visual gun-sight in an aircraft, it is essential to ensure that the range is measured on the correct target. This may be achieved by correlating the range and elevation, and the beam-switching technique (cf. § 5·3) in elevation may be used for this purpose. Fig. 4·6

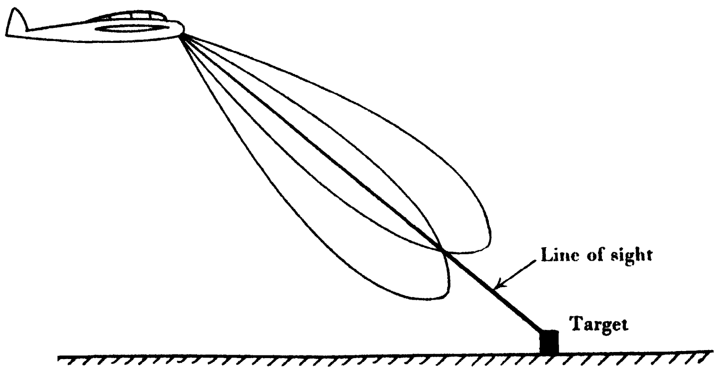


Fig. 4·6. Beam-switching for range-elevation correlation.

illustrates the system. The equisignal path of the beam-switching system is arranged to correspond with the line of the gun-sight, the echo intensities being displayed as deflexions on a cathode-ray tube. The target echo is characterized by being of the same intensity for the two beam positions.

It is desirable for this application to have some automatic indication of the range. To provide this a narrow strobe is used in such a way that its distance along the time base is decreased by the echoes from the 'up' beam and increased by the returns from the 'down'. If this is done, the range strobe will automatically stay at the intersection of the two beams and provide a continuous measure of the range of the object at which the sight is aiming.

Chapter 5

THE DETERMINATION OF AZIMUTH

Almost all radar equipments are concerned with measuring bearing or azimuth as well as range. Thus search radars, whether they are located on the ground, on the sea, or in the air require to determine both the range and the bearing of the target. In some cases the elevation of the target is also required. The methods used for bearing determination vary according to the application. In certain cases, e.g. when the radar is used for fire-control purposes, high accuracy is necessary, but usually when this is so information is required on only one target at a time. In others, and the various types of search radars come into this category, ability to present the general situation and track a number of targets simultaneously is required, and it is permissible to relax on the accuracy requirements to some extent.

Many of the methods used for the determination of azimuth have been developed from the methods of classical direction-finding.* This is particularly true of the longer-wave systems. In some cases the methods have undergone modification in detail if not in principle, as will be apparent from the following descriptions.

On the longest radar wave-lengths separate aerials are used for sending and receiving. In general the sending aerial is used to 'floodlight' a wide area in front of the station, and the receiving aerial is used for direction-finding. In the case of the G.L.† equipments ($\lambda = 3\frac{1}{2}$–$5\frac{1}{2}$ m.) a modified Adcock system is used for the latter purpose, whilst in the case of the C.H. stations ($\lambda = 7$–12 m.) crossed dipoles and a goniometer are employed. These methods are fully described in §§ 5·1·1 and 5·1·2, but it may be noted here that for these wave-lengths the dimensions of the aerials are such as to make it impracticable to use concentrated beams of radiation. The bearing is therefore obtained from the 'null' rather than from the setting for maximum signal.

* Most of these methods used vertically polarized aerials. Horizontal polarization is almost invariably used for radar (cf. § 8·1).

† Cf. p. 116.

With the introduction of shorter wave-lengths, however, it has become possible to use beamed aerials and to realize what is, in effect, a radio searchlight. As the beam sweeps round it 'illuminates' a narrow sector of the sky; any aircraft in that sector scatters radiation back to the station and is detected. With this system there is an obvious advantage in using the same beamed aerial for sending and receiving, and a method for doing this has been devised (cf. § 5·2·1).

With the introduction of rotating beam systems came a new form of display—the plan position indicator (P.P.I.). This takes the form of a rotating radial time-base display, the direction of the rotating radius being at all times parallel to the direction of the aerial beam. The received pulses are displayed as brightness modulation of the cathode-ray tube spot, and so aircraft responses produce on this presentation a short arc of brightness in a position relative to the centre of the tube which corresponds directly to the distance and direction of the target from the aerial system. In some cases, including the 1·5 m. coastal (low-flying) reporting chain (C.H.L.), ground control of interception stations (G.C.I.) and fighter direction stations, it is found convenient to track the aircraft directly on a map drawn on the face of the cathode-ray tube. Although not capable of the highest accuracy in azimuth determination, the plan position indicator system is very widely adopted on account of its continuous presentation of the overall situation coupled with the possibility which it offers of plotting a large number of echoes without losing the general picture. Some of the circuit principles used for producing this type of presentation are given in § 5·2·2.

For certain purposes, e.g. fire control, the very highest accuracy attainable is required for azimuth determination, and in such cases the 'split beam' or beam-switching technique is employed. In this an array is used which can send out either of two beams displaced slightly in azimuth. The beams are sent in rapid succession, and the display system is arranged to show the relative received amplitudes. When the amplitude received from the two beams is the same, it is known that the bearing of the target is exactly half-way between the two axes (fig. 5·1). This method gives a much greater accuracy than the simple beam, because (i) we are operating on the steep side of the polar diagram instead of the maximum where the rate of change of signal strength with angle is small, and (ii) the display system shows the sign of the signal inequality and indicates

which way the aerial has to be moved to obtain a balance. Several practical systems of this type are described in § 5·3.

While in this chapter we are necessarily concerned mainly with the azimuthal polar diagram of the equipment, it must not be forgotten that the vertical polar diagram of an equipment for the determination of azimuth is most important. Thus in the early 12 m. radar the difficulty always was to obtain adequate range at low angles of elevation. To be completely satisfactory as an aircraft-reporting station a full coverage in elevation from 0 to 90° may be desirable.

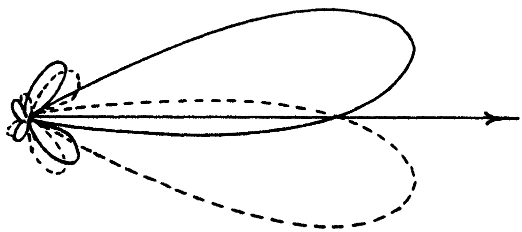


Fig. 5·1. Illustrating beam switching.

On the other hand, for surface-vessel detection a beam concentrated around 0° elevation may be preferred, while when working from a ship, the vertical polar diagram required will depend greatly on whether or not the aerial is mounted on a stabilized platform. These and other questions connected with the elevation polar diagram of radar direction-finding sets are discussed in § 5·4. Finally, in § 5·5 we give a short discussion of the non-instrumental errors which occur with these direction-finding systems.

5·1. Flood-lighting systems

5·1·1. *Crossed dipoles*

The principle of the crossed dipoles method is made clear in fig. 5·2. The dipoles are set exactly at right angles in the horizontal or x-y plane, one, the x-dipole, having its axis parallel to the line of shoot of the sender system. Suppose that the target lies in the direction OT from the station, the returned wave will induce e.m.f.'s with amplitudes*

$$E_1 = \frac{E_0 \cos[\frac{1}{2}\pi \sin\theta]}{\cos\theta} \quad \text{and} \quad E_2 = \frac{E_0 \cos[\frac{1}{2}\pi \cos\theta]}{\sin\theta}$$

* Half-wave aerials are used.

in the x- and y-dipoles respectively, and hence if the ratio of these two amplitudes can be measured, θ can be found. A goniometer is used for the ratio determination. This consists of two coils set at right angles, one coil being connected by a transmission line to the x-dipole and the other in the same way to the y-dipole. A third coil,

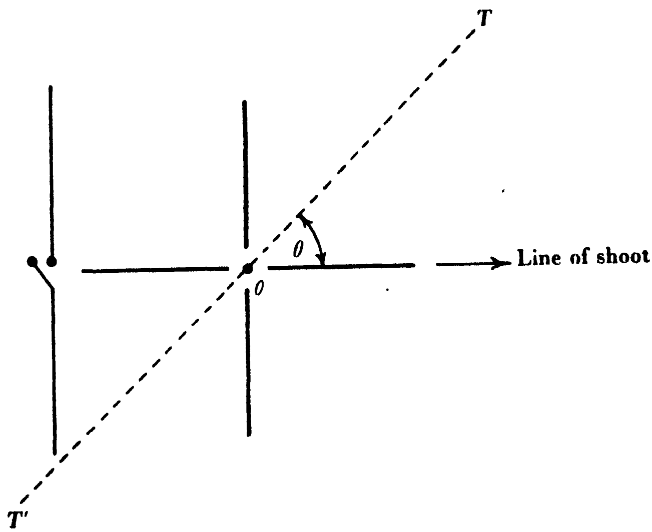


Fig. 5·2. Crossed dipoles for C.H. azimuth determination.

known as a search coil, is free to rotate between the two fixed coils, and it will be clear by reference to fig. 5·3 that the position (ϕ) of the search coil for zero current is given by

$$I_1 \cos\phi - I_2 \sin\phi = 0, \tag{5·1}$$

where I_1 and I_2 are the currents in the fixed coils (assumed to be in phase). If the transmission lines are electrically equivalent

$$|I_1/I_2| = |E_1/E_2| = \frac{\cos[\frac{1}{2}\pi\sin\theta]}{\cos\theta} \Big/ \frac{\cos[\frac{1}{2}\pi\cos\theta]}{\sin\theta} = \tan\theta \text{ approx.},$$

and hence $\tan\theta \approx \tan\phi$, so that the position of the search coil (ϕ) for zero current gives a measure (θ) of the azimuth required.

There is an ambiguity in this method of direction-finding, since if the signal arrives in the direction $T'O$, which differs by 180° from the direction TO, the setting will be the same. In order to resolve

this ambiguity, a reflector is placed behind the crossed dipoles. This reflector consists of two $\frac{1}{2}\lambda$ lengths of conductor which can be connected together by relay contacts. If the radar response is due to an aircraft in front of the station, the amplitude of the echo will increase on closing the relay, while if it comes from an aircraft behind the station the amplitude will decrease in similar circumstances. This method of sense-finding is also used in a number of the later systems.

The disadvantage of these so-called C.H. stations were realized even before they were built. Radiation leaks back inland producing

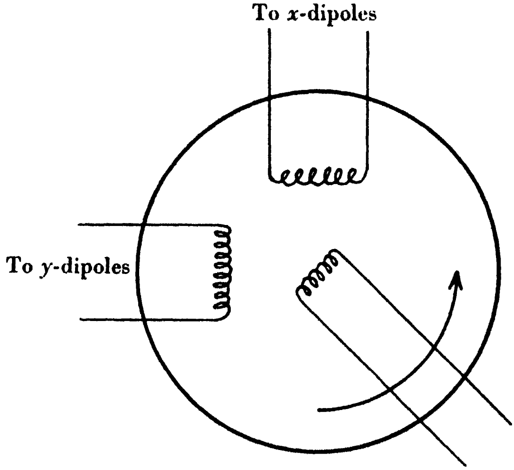


Fig. 5·3. Illustrating the principle of the goniometer.

echoes from hills and aircraft outside the zone of interest. Crossed horizontal dipoles are liable to polarization errors in direction-finding, and frequent calibrations to check the system are needed.

If the two transmission lines are not electrically equivalent the two currents (I_1 and I_2) are not in phase and a setting for zero current in the search coil is impossible. The formula $I_1/I_2 = \tan\phi$ is still approximately true, however, although the setting is now for minimum current and the quality of the direction-finding is degraded. If, in addition, a mismatch exists, serious quadrantal errors are introduced as well. The total feeder run between aerials and receiver at a C.H. station is often between 500 and 1000 ft., and the standard of workmanship to ensure exact electrical equi-

valence of the two lines is usually impossibly high. It is general practice, therefore, to include in each line a suitable phase-shifting network. This is used for the initial balancing of the system.

5·1·2. *Modified Adcock system*

This consists in using a rotatable array of two horizontal dipoles spaced one wave-length apart and fed in anti-phase. Such an array has a sharp null (cf. fig. 5·4), so that direction-finding on the null by

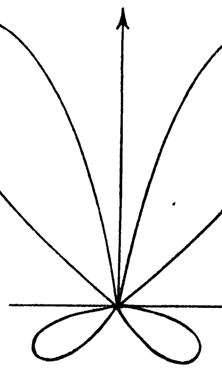


Fig. 5·4. Polar diagram of modified Adcock aerial system.

rotation of the array (instead of the goniometer) is possible. The modified Adcock method, in common with that using crossed dipoles, has the disadvantage that the signal disappears at the moment of taking the reading. This low signal condition is avoided in some of the later systems by adding the constant signal from a non-directional receiving aerial.

5·2. Rotating beam systems

The mean height of a C.H. aerial array is usually only about eight wave-lengths, and low-flying aircraft are not detected except at short ranges. For this reason additional stations were built to supplement the low-flying coverage of the C.H. stations. These stations (C.H.L.) operate on a wave-length of 1·5 m. They are sited on high cliffs, so that effective heights of the order of 50 wave-lengths are usual, and, in consequence, much improved performance on low-flying aircraft results. The C.H.L. stations have other advantages. The shorter wave-length allows the use of aerials

giving a narrow beam, and, as has been explained already, a display capable of much better performance in high traffic densities.

The aerials used for these C.H.L. stations are of the conventional broadside type (cf. fig. 3·1) consisting of four or five stacks of full-wave dipoles spaced 1·1–1·25 wave-lengths apart with generally four elements in each vertical stack. The beam in the horizontal plane therefore has a width of $\pm 11°$ to zero, and in the vertical plane the corresponding width is $\pm 30°$. A sheet of wire-netting (usually of $\frac{1}{2}$ in. mesh) is employed as an aperiodic reflector, and this is spaced $\frac{1}{8}\lambda$ behind the array, this spacing having been found to give the best compromise between maximum forward gain and minimum back radiation.

In addition to the 1·5 m. stations much use has been made of search radars operating on still higher frequencies, particularly on 3000 and 10,000 Mc./sec. For ground and ship radars the higher frequencies offer the possibility of improved low-flying performance, as well as the advantages of P.P.I. presentation with a higher definition (due to the narrower beam widths attainable) and give greater accuracy in direction-finding. The higher frequencies also allow the use of relatively narrow-beamed radars even for airborne sets where there is an obvious restriction on the size of the aerial which may be used. Thus an A.S.V. (air to surface vessels) radar may be designed with satisfactory P.P.I. presentation with only a 3 ft. aerial (a permissible size for aircraft fitting). At a frequency of 3000 Mc./sec. this means a beam width of $\pm 6°$.

5·2·1. *Common aerial systems*

If in a beamed system we can use a single aerial for both sending and receiving, we can economize in space and equipment and avoid all the problems of synchronous rotation of two aerials. We can also satisfy the requirement that the sender shall 'illuminate' the same portion of space that is 'viewed' by the receiver. A method of doing this for frequencies in the 200–600 Mc./sec. region was devised during the war. It is illustrated schematically in fig. 5·5. The action of the system is as follows. During the period of each sender pulse both the spark gaps (X and Y) break down and become equivalent to r.f. short circuits, so that infinite impedances are presented at A and B respectively, and the sender power passes direct to the low-impedance aerial instead of entering the receiver. Immediately the

sender pulse ceases, the spark gaps cease to conduct and a short-circuit appears at A and is transformed to look like an infinite impedance at B. All the incoming energy therefore flows past Y into the receiver.

The same principles are employed in using a single aerial for sending and receiving on the centimetre wave-lengths. There are, however, important differences in design, low-pressure gas-filled

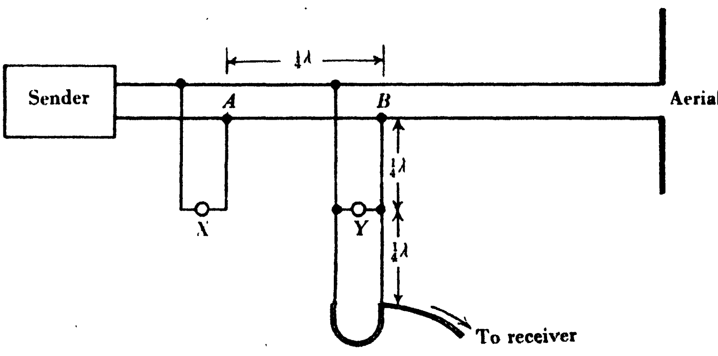


Fig. 5·5. Common aerial working (X and Y are gas-enclosed spark gaps).

resonant cavities* being used as breakdown elements instead of spark gaps. In all cases, however, the spark gaps are required to ionize and de-ionize rapidly, and this involves enclosing the gap in a glass envelope and filling at reduced pressure.

5·2·2. *Plan position indicators*

In the P.P.I. display, which is used with a rotating beamed aerial, the signal appears as an intensification of the cathode-ray tube spot and the time base is radial from the tube centre. The time-base deflexion is usually obtained magnetically, suitable current coils being placed around the base of the tube. A current having a linearly rising saw-tooth wave form is passed through these coils, and since it is important that the current shall be zero at the beginning of the trace a balanced wave form is usually used, a negative current passing in the interval between traces when the spot is blacked out. In ground stations the current coils are generally mounted around the

* Cooke, A. H., 'Gas Discharge Switches for Single Aerial Working', *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 186 (1946).

neck of the cathode-ray tube so that they can be rotated, a synchronous motor arrangement being used to repeat the azimuth of the aerials at these coils. An alternative arrangement, more usual for airborne equipments, is to use two fixed coils at right angles on the neck of the cathode-ray tube, and feed these through a goniometer-like arrangement acting as a current transformer. The saw-tooth current wave form may thus be fed into the rotor winding and the cathode-ray tube coils fed from two stator coils at right angles. In this way the azimuth of the time base on the tube will repeat the azimuth of the goniometer device, which being directly geared to the aerial shaft repeats in turn the azimuth of the aerial.

The rectified radar signals are applied to the grid of the cathode-ray tube so as to intensify the spot, but since too great a voltage on the grid causes de-focusing, a signal limiter is necessary. This is usually set to limit at a level of about twice the noise level, so that signals of any size greater than this give a full spot intensification, while the noise peaks themselves are just visible 'splashes'. A long afterglow tube is, of course, used, so that the complete picture is present at any time—the afterglow duration being chosen to be just greater than the normal period of revolution of the aerial. The actual echoes appear as arcs of brightness, and to obtain the azimuth, or the actual target position, the centre of these arcs has to be estimated. Plate 1 shows a photograph of a typical P.P.I. picture.

For some purposes plotting is done in polar (r, θ) coordinates, and it is then convenient to display the range and azimuth calibrations of the tube as faint bright lines on which the echoes appear superimposed. This can be done electronically, suitable pulses being generated in an auxiliary circuit (some mechanical azimuth switch being usually required) and mixed with the radar signals. For other purposes (a good example being the close control of fighter aircraft in interception operations), a magnified picture is required, and this result is achieved by providing around the neck of the tube an additional fixed-coil magnetic system which is employed to deflect the centre of rotation of the cathode-ray tube spot to any desired point outside the periphery of the tube. The 'scan' is then expanded and any selected portion suitably magnified.

In most cases a linear time base is required, as already described, although sometimes a time base whose speed decreases slightly with increasing range is acceptable. The latter is easier to generate, being

the initial part of an exponential curve, and the slight distortion of the resulting map grid may not be serious. In a few cases markedly non-linear time bases are required. For example, in H_2S (more fully considered in § 5·4 below) the slant ranges actually obtained from the aircraft have to be interpreted as ground ranges in order that the P.P.I. shall present a map of the country over which the aircraft is flying. From fig. 5·6 it will be apparent that the time-base sweep must not start until a time $2h/c$ after the sender pulse, for this is the time delay of the first ground return. Thereafter the sweep

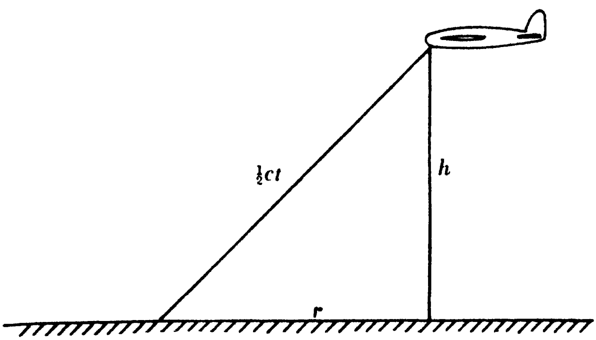


Fig. 5·6. Illustrating H_2S principle.

must move rapidly at first and then more slowly, following the law $r = \sqrt{[(\frac{1}{2}ct)^2 - h^2]}$. Circuits for doing this have been devised, but we must refer the reader elsewhere for details.*

5·2·3. *Range-azimuth presentation*

An alternative to the P.P.I. which is sometimes used when plotting in polar coordinates is a presentation in which echoes are displayed on a cartesian plot of range and azimuth. The signals are applied as intensity modulation as for the P.P.I., but the time base is a vertical line which is moved across the screen to correspond to changes of aerial azimuth. Echoes therefore paint on the afterglow screen as horizontal lines, whose centre indicates the azimuth of the echo, and whose length depends on the beam width of the radar. The range-azimuth type of presentation is much used for airborne radars, in which case the zero-azimuth direction, displayed in the

* Williams, F. C., op. cit.

centre of the c.r.t. screen, corresponds to the direction of the aircraft heading, and a reading to the right or to the left indicates the degrees off-bearing.

5·3. Beam switching

The beam-switching system is widely used for accurate direction-finding, but the principle of the method is the same whether it is carried out when sending, when receiving, or both. Fig. 5·7 illustrates the system used in the case of the early 1·5 m. coast-watching radars. The switching (received signal only) is carried out

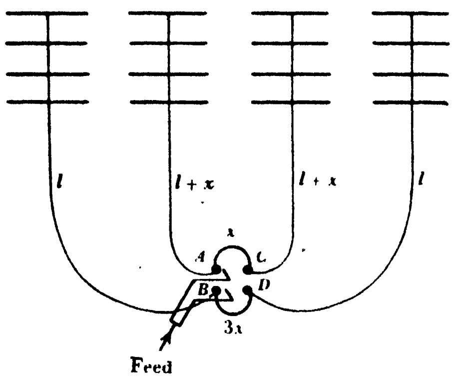


Fig. 5·7. Beam switching.

at about 20 c./sec. using a rotating mechanical switch which provides two alternative paths from the four bays of the array to the receiver. With terminals A and B connected the feeder lengths are respectively l, $l+x$, $l+2x$, and $l+3x$, while with C and D connected the lengths read as above but in the reverse order. The length x is chosen so as to give an angle between the two beam maxima of about 11°, enabling azimuths to be measured to something better than $\pm\frac{1}{4}°$. This method has now been largely superseded by a method using electronic switches. The principle of reversing the phase distribution across the aperture of the array is the same, but the switching is accomplished by means of diodes which are made alternately conducting and non-conducting (fig. 5·8).

The display is of the range-amplitude type with the trace slightly displaced laterally in synchronism with the beam switching. The signals corresponding to the two beam positions then appear side

by side, and are adjusted to equality by rotating the whole aerial system, giving a very accurate azimuth setting. In some cases a stroboscopic disk containing red and green filters (fig. 5·9) is rotated in front of the tube, and the two echoes then appear as different colours and matching the two signal amplitudes is rendered slightly easier. In other cases, and this applies particularly to air-borne radars operating on the metre wave-lengths, the back-to-back

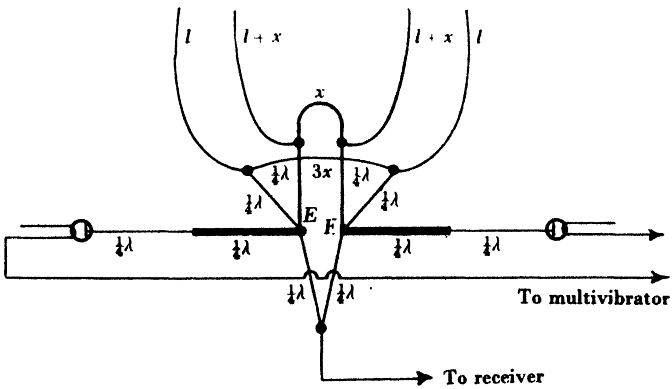


Fig. 5·8. Alternative arrangement for electrical beam switching.

display shown in fig. 5·10 is used. The 'split' signals are then displayed on either side of the trace and equality results in the trace bisecting the signal line, indicating that the target is dead ahead. An off-bearing target is indicated by the centre of the composite signal displaced to the right or left, i.e. in the same direction as the target.

Since angle measurement with the beam-switching method involves the matching of two signal amplitudes, rapid fading of the received signal introduces a strong disturbing element, far more so, in fact, than with range measurement. This is particularly serious if for any reason the fading of the echo has a marked periodicity at the beam-switching frequency. Moreover, side lobes in the aerial diagram may give rise to ambiguities since there may be two or more directions for which equal signals can be obtained (cf. fig. 5·1). Where this is likely it is usual to feed the side elements of the array less strongly than the centre elements. This gives a slightly wider main lobe than with uniform feeding, but the side lobes are smaller (cf. §3·1·1).

The beam-switching technique is also widely used in centimetric radars for accurate direction-finding. Where both elevation and azimuth information is required, it is convenient to use a para-

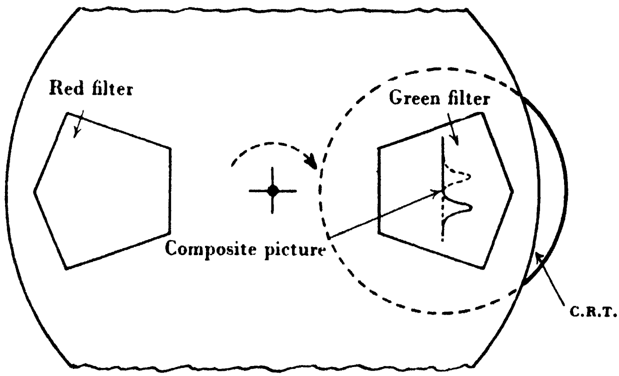


Fig. 5·9. Stroboscopic disk arrangement.

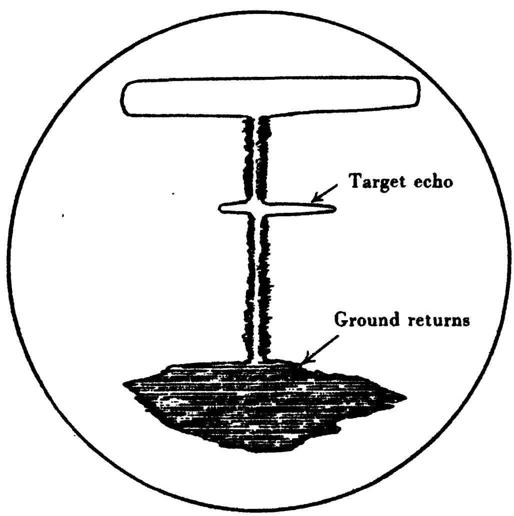


Fig. 5·10. Back-to-back display used with metric radars.

boloidal mirror system fed by a single dipole, since if the dipole is displaced slightly from the centre of the mirror, the axis of the beam will be displaced an equivalent amount in the opposite direction.

Hence, if the dipole is mounted eccentrically and spun about the axis of the mirror, the axis of the beam will describe a cone about this line. It can be arranged that (i) a single pulse is sent out by the sender at each of the dipole positions corresponding to the 3, 6, 9 and 12 o'clock positions, or (ii) that a number of pulses are sent out in each quadrant centred on these clock positions, the angle over which the pulses are sent out being, say, 45°. In either case, by comparison of the signal amplitudes received respectively from the left and right positions of the beam, and similarly by comparison of the amplitudes from the upper and lower positions, we may centre the beam on the target and hence determine both azimuth and elevation.

All this has been described as though only the sender dipole were eccentric; it is, however, immaterial whether the sender or receiver or both dipoles are eccentric; the effect can be the same. It is generally convenient to arrange a small eccentricity of both the receiver and sender dipoles rather than the larger eccentricity which would be required if only one were eccentric. It should be noted that since the plane of polarization rotates with the dipole, it is essential that both dipoles should rotate in synchronism, whether both are eccentric or not.

When azimuth information only is required, it is often convenient to use two identical dipoles each displaced on either side of the focus of the paraboloid, and switch these to the receiver in turn. The dipole displacement is chosen so that the beams overlap, and so that the cross-over position may be used for accurate direction-finding as already described. Using two aerial elements in this way the mechanical complications of a moving source are avoided. On the other hand, cross-talk effects may occur with this system and may be serious.

5·4. The importance of the vertical polar diagram

We have already noted that for ground or shipborne search radars complete viewing from 0 to 90° in elevation is desirable. However, it is usually unnecessary to be able to detect aircraft above, say, 50,000 ft. in height, so that the ideal coverage is roughly as shown in fig. 5·11, from which we note that the gain of the aerial should fall off with increasing elevation (α) above the value (α_m) corresponding to 50,000 ft. at maximum range. Neglecting earth's curvature, it is

easy to see that the vertical polar diagram should follow the law $\phi(\alpha) \propto \operatorname{cosec}\alpha$ for $\alpha > \alpha_m$, $\phi = \text{constant}$ for $0 < \alpha < \alpha_m$ and $\phi = 0$ for $\alpha < 0$, remembering (cf. eqn. (3·13)) that the range of the radar is proportional to the field-strength factor ϕ of the aerial. If this diagram can be achieved a constant response independent of distance is provided from an aircraft flying at a fixed height once $\alpha > \alpha_m$.

These ideal requirements are difficult to realize with long-wave radars, particularly the sharp cut-off near $\alpha = 0$, which is essential to avoid the complications of earth's reflexions, interference gaps, etc. The attainment of this obviously involves the use of an aerial with large aperture/λ ratio. Generally for frequencies less than about 600 Mc./sec., the aerials cannot be designed sufficiently directive to be independent of earth reflexions and so interference gaps result.

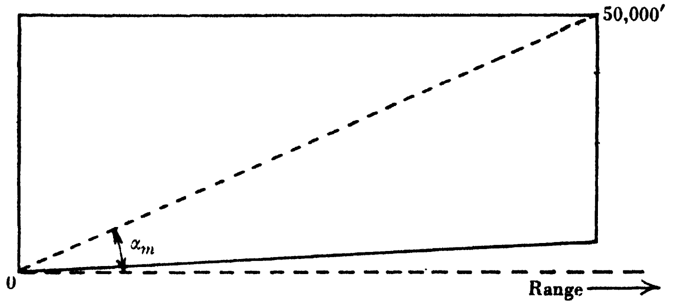
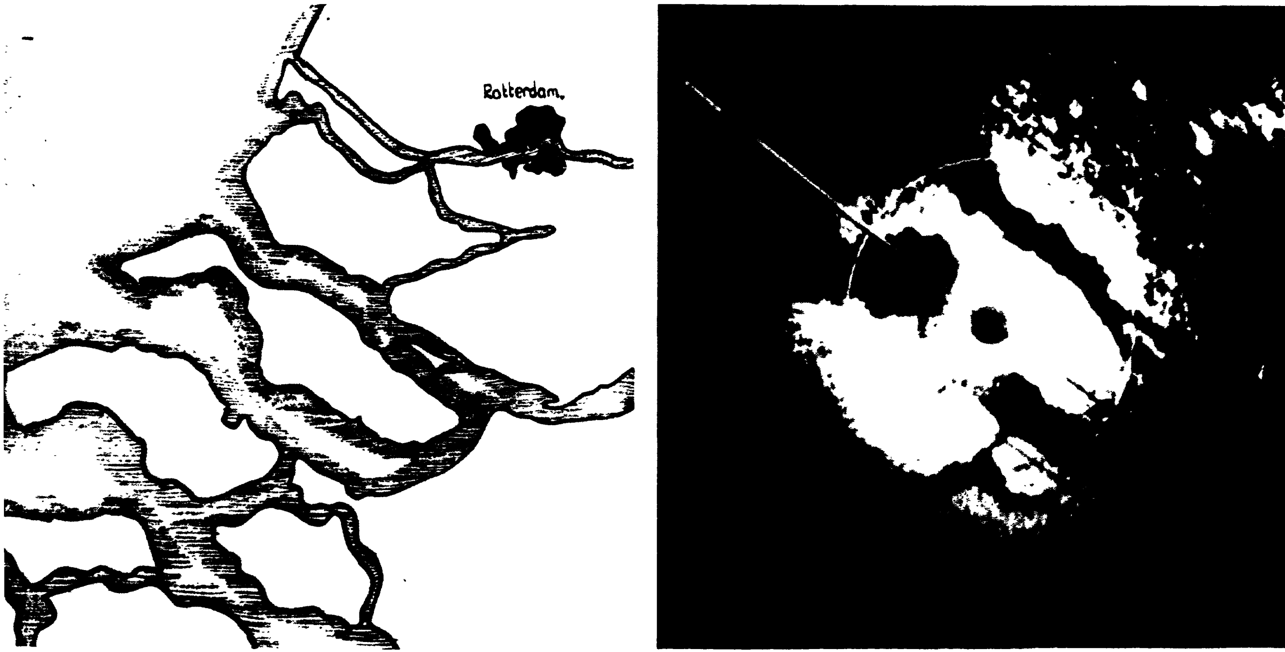


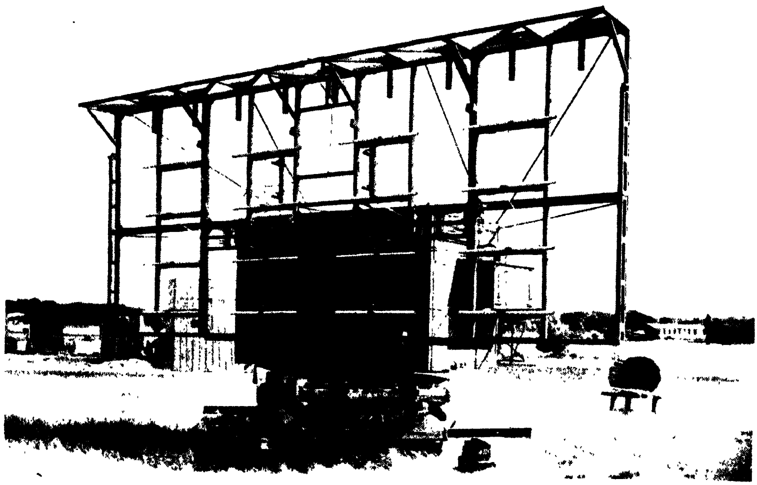
Fig. 5·11. Ideal V.P.D.

Various methods of providing something approaching these requirements are possible, the simplest being the use of an auxiliary aerial which is arranged at a different height above ground from the main array. The heights are arranged so that many of the gaps in one vertical polar diagram are filled by the lobes of the other and, by switching as required (either manually or continuously and automatically) from one aerial to the other, a target may be followed without interruption.

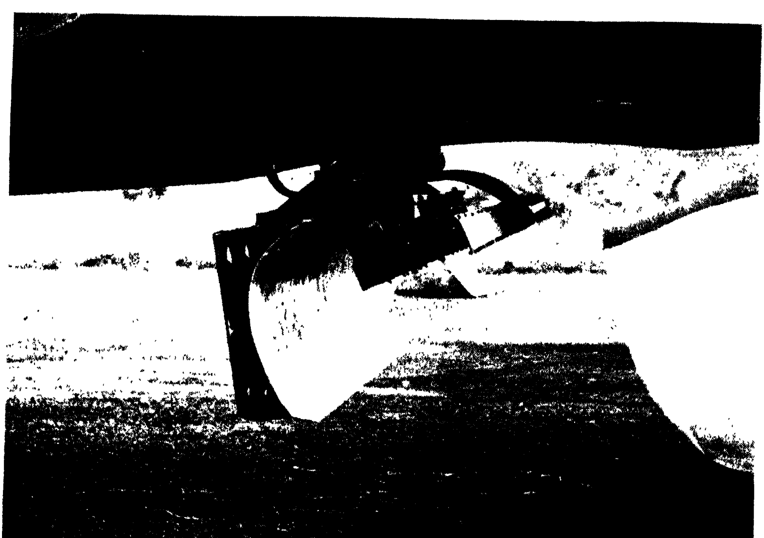
Another method when the reflexions occur from land as distinct from water is to use vertical polarization, in which case gaps at elevations above about 8° are much reduced since the reflexion coefficient (cf. fig. 3·5) becomes small, although, of course, the range in the lobe maxima is correspondingly reduced. This method is, however, useless when there are many gaps at low elevations, as are



Example of radar map obtained with H_2S.



(*a*) Broadside array of mobile Ground Control of Interception (G.C.I.) station (cf. § 6·3). The upper row of vertical radiators is part of the identification system (cf. § 10·3).



(*b*) H_2S aerial installation with nacelle (part of which is showing on right) removed. The reflector profile is designed to produce a cosecant law (cf. § 5·4). The feed is a wave guide with slot radiators.

obtained from a high aerial. Still another method of gap-filling which is useful on the longer wave-lengths is the use of an unsymmetrical array, or a symmetrical one tilted upwards so as to reduce the intensity of the reflected ray. The gap-filling obtained by this method varies with elevation, and with the aerial apertures (vertical dimensions) usually available on the longer wave-lengths is generally such that the first interference gap is not well filled.

On the centimetre wave-lengths, sufficient directivity is obtainable to render the vertical polar diagrams substantially independent of the earth reflexions. Furthermore, the radiation pattern of the array is usually much easier to control. Thus if the aerial consists of a flared wave guide feeding into a reflector the radiation pattern can be adjusted to give the required cosecant law by suitably shaping the profile of the mirror. The exact profile depends on the value of α_m for the equipment concerned, but a vertical section through the mirror generally approximates to the quadrant of a circle joined to a parabolic arc, while sections containing the normal to the plane of the 'fan' beam are parabolic. The production of the required reflector is usually quite difficult, and largely because of this the more modern equipments produce the required beam shape in two stages. A primary radiator consisting of a linear array* or 'cheese' is used to produce the necessary beaming in azimuth, and this is then employed to feed a reflector which produces the required elevation pattern.

Cosecant aerials are also used for some airborne applications. The two most important cases are H_2S and A.S.V. The former has already been referred to in connexion with fig. 5·6. It is an airborne navigation and blind-bombing radar aid whose action depends on the fact that the signals returned from the earth's surface vary in magnitude according to the characteristics of the terrain. In particular, the responses from water surfaces are found to be relatively weak, those from open country much stronger and those from built-up areas much stronger still. A P.P.I. with the non-linear time base already mentioned is used, so that a map of the country over which the aircraft is flying appears on the screen. Plate 2 shows a typical H_2S P.P.I. screen; the correlation with the principal land features should be noted.

For H_2S, and to a less extent for A.S.V., the use of the correct cosecant law is important to ensure that the strength of the signal

* Ratcliffe, J. A., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 22 (1946).

depends only on the terrain and on the aircraft height, not on the range. The case is exactly analogous to the property of this aerial already mentioned of giving constant signal strength on an aircraft flying at constant height, except that the aerial has to be mounted the other way up. The aerials for both H_2S and A.S.V. are required to produce a beam narrow in the horizontal plane, and they must therefore have considerable extent in that plane; further, they must be stabilized against pitch and roll.

Stabilization is also important in the case of ship radars. When the measurement of elevation is not required and when a large aerial gain is not necessary, aerials for ship radars are generally designed with broad vertical beams to eliminate loss of target through pitch and roll. Such arrays, when rotating continuously, are not stabilized against yaw and turning of the ship; instead, the display is stabilized to give bearings relative to true north. The direction-finding readings are then, however, subject to cross-roll and deck-tilt errors, and these errors may be serious for elevated targets. In consequence, in the more recent ship radar installations, the radar aerial itself is stabilized so that its reference axes have a fixed orientation in space.

5·5. Direction-finding errors due to site effects and atmospheric causes

In addition to errors in direction-finding arising as a result of instrumental deficiencies, errors may occur due to (i) non-rectilinear propagation due to atmospheric causes and (ii) site effects. Errors due to atmospheric causes are usually very small, such measurements as have been made indicating that the error seldom exceeds a few minutes of arc. On rare occasions, such as when making a measurement which involves a wave which has traversed a path along the leading edge of a warm front, larger errors are encountered. Errors due to site effects are usually far more serious, but these can be predicted from the simple theory considering the terrain within a few hundred feet of the radar station. They are usually most serious when one is working near a gap in the vertical polar diagram of the aerial system concerned (including the effects of any earth reflexions). It is convenient to differentiate between the effects due to reflexions from the earth's surface, and those due to diffraction by obstacles in the path of the transmission. Provided the land is reasonably flat and there are no obstacles of any appreciable size near the radar

aerial, absolute errors in direction-finding on metric wave-lengths (C.H.) should rarely exceed 1 or 2°, and many of these should be capable of elimination by calibration. With beam systems the figure is much less. On centimetric wave-lengths errors due to the ground having a transverse slope should be very small, although errors due to diffraction effects when viewing targets from sites with prominent local screening may be as high as 1 or 2°. It is very important therefore to choose a site for a radar station with care, and the rule is to choose a site which allows a clear optical view of the targets which it is hoped to 'see', and avoid sites with appreciable transverse slopes or local screening.

Chapter 6

THE DETERMINATION OF ELEVATION

We have seen in the preceding chapter that most radars determine, in addition to the range, the azimuthal angular coordinate of the target. The remaining coordinate is the elevation angle, methods for the determination of which are considered in this chapter. In certain equipments, however, the determination of the two angular coordinates is so interdependent that the problem is considered separately in Chapter 7.

When targets on water or land are being detected from the air (as in A.S.V. and H_2S) the elevation coordinate varies but does not need to be separately determined, since it can be deduced from the range when the aircraft height is known. This has already been fully considered in Chapter 5, where it was seen that normal azimuth-determination technique is used with the proviso that the fan beam must also be suitably shaped in elevation. Special facilities for elevation-finding are important for aircraft reporting (and, in wartime, for interception and fighter control) by ground and ship-borne stations, and also for air-to-air systems such as the wartime A.I. Some forms of A.I. make use of two-dimensional scanning, which is considered in Chapter 7. The methods to be considered in the present chapter are those used by the ground and shipborne stations which determine elevation, but do not employ two-dimensional scanning.

6·1. Ground and shipborne radars for aircraft detection

The ground and shipborne radars designed for locating aircraft may be grouped into three classes, viz. those determining range and azimuth only, those determining range and elevation only, and those determining all three coordinates. The first class are generally beam stations, or flood-lighting stations whose site is unsuitable for elevation-finding, and they will not be considered further here. The second class are supplementary equipments used in conjunction with some other azimuth-determining station. The third class exhibit the greatest variety of design—they may be flood-lighting stations, pencil-beam stations, or stations beamed in azimuth but

determining elevation by steerable zero or signal-comparison methods; some may be continuously rotating to give uninterrupted azimuth indications, others may be directed manually on to a particular target.

6·2. Flood-lighting stations

The flood-lighting technique has already been described (cf. § 5·1) in relation to the measurement of range and bearing. The determination of elevation by this class of station involves using the interference effects between the direct wave from the target to the radar receiver and the wave which reaches the receiver after reflexion at the ground. If the reflexion takes place from substantially flat ground the interference diagram is calculable (cf. § 3·1·2), and this provides perhaps the simplest method of elevation-finding. No special apparatus is necessary, and the method consists in noting the initial 'pick-up' range of the target and the variation of received signal strength as it approaches. A comparison of the results obtained with the known performance diagram of the station provides information about the elevation of the target, but in many cases the results are difficult to interpret, and if the aircraft is changing height or aspect rapidly the method may be misleading.

6·2·1. *Spaced aerial method*

A method which avoids many of these difficulties is the so-called spaced aerial method. This is an adaptation of methods worked out by Friis* and Wilkins† in connexion with ionospheric studies. In its first form it consisted in measuring the phase difference between the e.m.f.'s induced in two similar and parallel horizontal aerials at the same height (h) above ground, but it was the later variation of the method involving the comparison of the e.m.f.'s induced in two aerials at different heights which found application in radar.

Following the treatment of interference effects already given (cf. § 3·1·2), it will be noted that if we consider a receiving aerial at a height (h) above ground (assumed flat and with reflexion coefficient $\rho = -1$), the direct wave reaches the aerial with a phase advancement (compared with the wavelet from O, fig. 3·6) given by

* Friis, H. T., *Proc. Inst. Radio Engrs, N.Y.*, vol. 16, p. 658, 1928.
† Wilkins, A. F., *J. Instn Elect. Engrs*, vol. 74, p. 582, 1934.

$\phi = hk\sin\alpha$, where $k = 2\pi/\lambda$, and the reflected wave arrives with a phase retardation of the same amount, and, in addition, a phase shift of 180°. Therefore the amplitude and phase of the signal received directly may be represented by $f(\alpha)\,e^{jhk\sin\alpha}$, and the signal received after reflexion at the ground by $-f(\alpha)\,e^{-jhk\sin\alpha}$, where $f(\alpha)$ is the function corresponding to the 'free space', vertical polar diagram of the aerial.

The resultant signal is therefore

$$E = 2jf(\alpha)\sin(hk\sin\alpha), \tag{6·1}$$

and it has a phase which is always in quadrature with the phase of the wavelet from O.

If we use two receiving aerials at heights h_1 and h_2 above the ground as described above, the ratio of the two signal amplitudes is evidently

$$\frac{\sin(h_2 k\sin\alpha)}{\sin(h_1 k\sin\alpha)}.$$

Hence if h_1 and h_2 are known, the measurement of the amplitude ratio enables α to be measured.

This is the system which is used for C.H. and for the long-wave G.L. ('gun-laying') sets. As an example consider the C.H. medium-elevation height system. Here we have $h_1 = 2\lambda$ and $h_2 = 6\lambda$, so that for small angles of elevation the induced e.m.f.'s vary with angle of incidence, as shown in fig. 6·1 (*a*). The ratio of the e.m.f.'s is shown in fig. 6·1 (*b*), and it is clear from this curve that a knowledge of the ratio determines the angle of elevation uniquely up to 7·5°. In both the C.H. and the G.L. systems a goniometer is used for the signal comparison. This is satisfactory, since the e.m.f.'s induced in the two aerials are both in quadrature with the phase of the wavelet from O (fig. 3·6), and therefore they are in phase (or in anti-phase) with each other. If, therefore, the two aerials are connected to the field coils of a goniometer by two feeders of equal length, the currents in the two field coils will also be in phase. Further, if the search coil of the goniometer is set at an angle ϕ to the field coil connected to the upper aerial, zero signal will be obtained when

$$\tan\phi = \frac{\sin(h_2 k\sin\alpha)}{\sin(h_1 k\sin\alpha)}. \tag{6·2}$$

This is the simple theory. In practice, the results so obtained must

be modified somewhat, due to site variations, but the errors are usually small and can be allowed for by flight calibration.

One unavoidable feature of the method is that it gives ambiguous heights*—i.e. there may be several heights of aircraft at a given range which produce the same electrical effect in the receiver. These

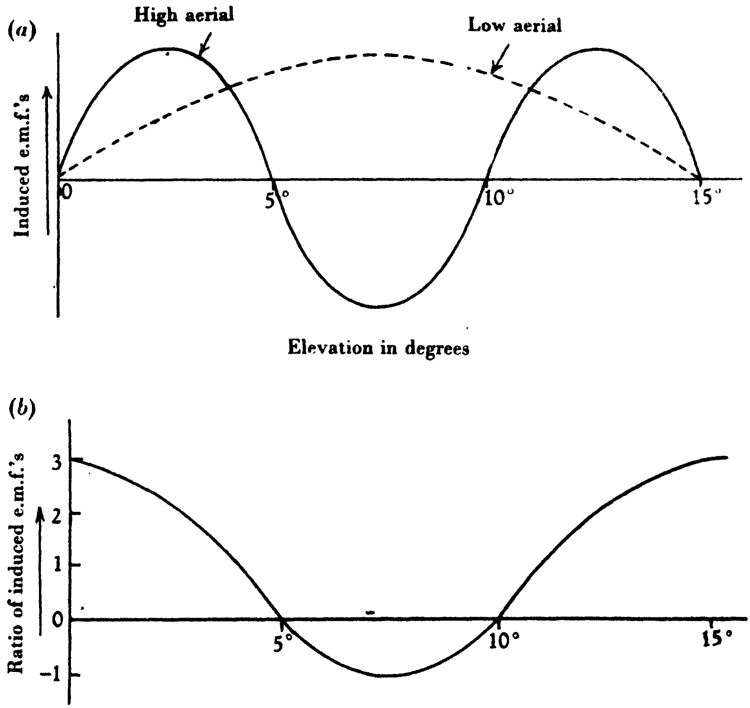


Fig. 6·1. Elevation-finding with aerials at different heights.

heights, in general, differ very considerably and may be resolved by intelligent 'filtering', as one will tend to change rapidly as the target approaches, while the other will tend to remain more nearly constant. Alternatively a third height-finding aerial may be used.

6·2·2. *Effect of site*

The chief disadvantage of the spaced aerial method of height-finding is that it requires a flat or gently sloping site in order to be

* The height (H ft.) is obtained from the range (r in miles) and the angle of elevation (α), using $H = 5280\ (r \sin \alpha + r^2/2\cdot 5R)$, where R miles is the radius of the earth. This calculation is often performed by mechanical means.

effective. Any considerable land irregularities near the station may render the system inaccurate or impossible to use. On many of the C.H. stations using this method, elevation-finding has to be restricted to certain regions in azimuth over which the land is reasonably flat. Within its limitations, however, it may be said that the method works well and, with few exceptions, stations give the expected accuracy of $\pm\frac{1}{4}°$ for aircraft flying above about 5000 ft. Considerably greater inaccuracies arise if the system is used outside the limitations imposed by the site. Thus one unavoidable feature of the spaced aerial method of height-finding is that it becomes inaccurate at small angles of elevation, and a lower limit must be set for each site. On average C.H. sites this is about $1\frac{1}{2}°$. Flight calibrations are also generally essential.

6·3. Stations beamed in azimuth—G.C.I.

The application of the spaced aerial method of elevation-finding to the beamed search radar stations is relatively straightforward, providing that it is possible to aline the aerial beam in the direction of the target and maintain this setting during the time that the goniometer (or its equivalent) is adjusted for a zero reading. For many purposes, however, this interruption of the plan position display is unacceptable, and a method is then needed which allows elevation-finding with a continuously rotating aerial. A number of suggestions for doing this have been put forward, but the most successful, and the one which has been widely adopted for the G.C.I. stations, uses two receiving aerials at different heights above ground. The two received signals are presented side by side as deflexions on a cathode-ray tube in a similar way to that already described for 'beam switching', a diode switch being used to connect the two aerials to the receiver in rapid succession, just as in azimuth determination (cf. § 5·3). A long afterglow tube is used to 'hold' the transient signals, and although the time the target is 'in beam' is short, the signals are retained in afterglow sufficiently long to enable an experienced operator to estimate the ratio* of the two signal amplitudes. This estimation is usually done by eye, no mechanical or optical aid being employed. This may be thought to be rather surprising, but during the war R.A.F. operators, who were

* Since the ratio is estimated directly, it is necessary to ensure that the receiver signal circuits follow a linear law.

familiar with interpolating in tenths of a grid square on plotting maps, found they were able to estimate the ratio with an accuracy better than 5 % of the larger signal. This was sufficiently good, since the practice was to take several ratios on successive rotations of the aerial, to plot the results on a height-conversion board (fig. 6·2), and obtain an average value for the height of the target. The method was remarkably successful.

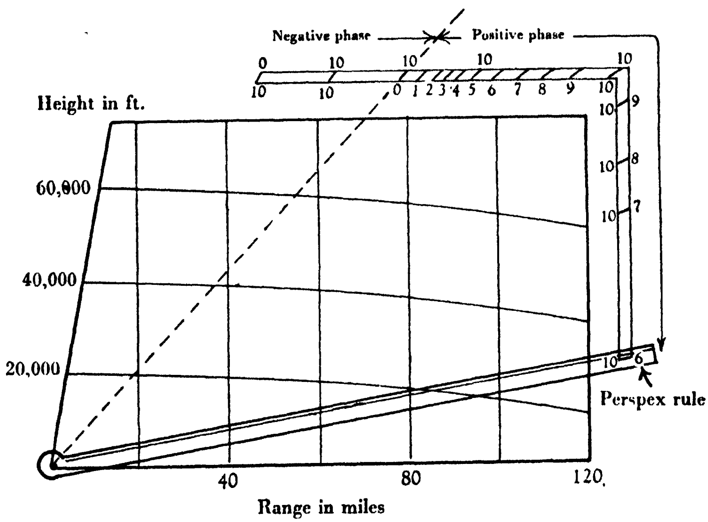


Fig. 6·2. Height-conversion board for G.C.I.

There are many different versions* of G.C.I. sets. We shall have to be content here in describing one particular set, choosing as our example the so-called mobile G.C.I. equipment. This used a 4-bay 4-stack broadside array (cf. fig. 3·1) with its centre 10 ft. ($= 2\lambda$ for the wave-length used) above ground, mounted on the side of a rotating trailer. For elevation-finding the array was divided into two halves, viz. the upper two stacks and the lower two stacks. Fig. 6·3 shows the vertical polar diagram given by such receiving aerials on the assumption that the ground is a perfect conductor and is perfectly flat. The same figure shows the ratio of the two signal strengths as a function of the angle of elevation, from which it will be seen that the system is insensitive to changes in elevation

* Taylor, D. and Westcott, C. H., *J. Instn Elect. Engrs*, vol. 93, part III A, p. 588 (1946).

below about 4° elevation, but that it allows satisfactory elevation-finding between 4 and 21° providing positive and negative phase can be distinguished.

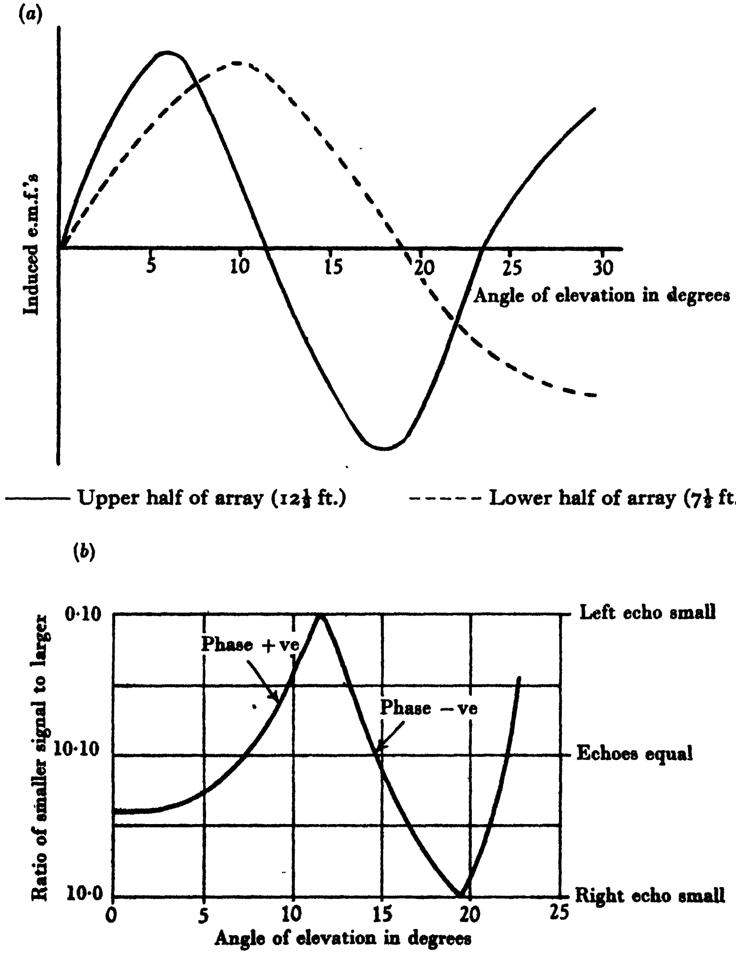


Fig. 6·3. Elevation-finding with divided array ($\lambda = 5$ ft.).

From fig. 6·3 it will be seen that the receiver is sensitive to signals from aircraft for all elevations up to at least 25°, but fig. 6·4 shows that the sender, if connected into the whole array in parallel, will give rise to a serious gap at 14° elevation. This can be overcome if

we switch the sender so that it feeds the two halves of the aerial in anti-phase. This is preferable to feeding into one-half of the array only, since it gives better gain at the maximum of the lobe; it also has the advantage that by noting whether a target is seen most clearly with the sender in phase or anti-phase, we obtain the 'phase check' which we desire for resolving height-finding ambiguities. It is desirable to be able to switch the sender from phase to anti-phase rapidly as is done for the receiver, but this was not possible with this equipment. Later a switch capable of handling the sender

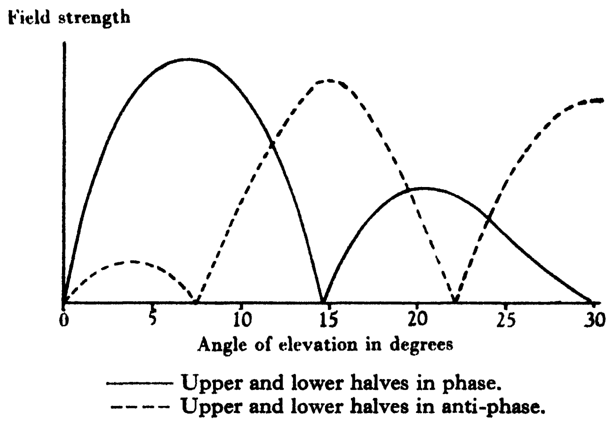


Fig. 6·4. Vertical polar diagrams of sender aerial.

power was developed and was incorporated in some of the later G.C.I. equipments. With the earlier G.C.I. radars the phase check could only be made by stopping the rotation of the aerial and comparing the signal strength with the phase switch in the two positions. In actual operation the phase check was not often required, since most targets were first seen below $11\frac{1}{2}°$ elevation.

6·4. Elevation-only stations

With the introduction of high-power centimetric radar the spaced aerial method of height-finding was largely abandoned and replaced by the variable elevation beam (V.E.B.) method. The use of a very narrow beam, controllable in elevation to determine height, renders the performance of the radar free from dependence on the surrounding terrain except only for the very low elevations for which

appreciable energy is reflected by the ground. The height-finding accuracy is therefore mainly dependent on controllable factors such as the beam width, the over-all mechanical accuracy and the method of display. The only uncontrollable source of error is that due to non-rectilinear propagation arising from refraction effects. This effect is fortunately small under normal circumstances.

The first V.E.B. system to be designed operated on a wave-length of 1·5 m. The aerial consisted of 56 full-wave dipoles in a vertical stack, and elevation-finding was carried out by locating the direction of maximum response as the beam was tilted up and down. The movement of the beam was achieved electrically by varying the relative phases of groups of dipoles.

There are many types of centimetric V.E.B. equipments, but the majority of the modern equipments use an aerial system consisting of a cylindrical parabolic reflector fed with a line source—usually a slotted wave guide. The vertical aperture is often quite large—for instance in one of the R.A.F. sets the vertical aperture is 60λ giving a beam width of about $1\frac{1}{2}°$. Motion of the beam is provided by mechanical or electrical means, the favourite electrical methods being by variation of the guide size, or by motion of a dielectric section in the guide. Both these methods vary the relative phases of the primary radiators (i.e. the slots on the guide) and hence swing the beam in elevation.

The beam has of course to be set in azimuth from the indications of the associated search radar equipment. The general practice is to use an azimuth marker which indicates on the P.P.I. of the search radar the orientation of the V.E.B. system in azimuth. If, in addition, a range marker is used, the P.P.I. operator can indicate at once to the V.E.B. operator the range and bearing of the target on which a 'height' is required, and know from the positions of the azimuth and range markers whether the correct target is being investigated (c.f. footnote, page 60).

To obtain very accurate elevation-finding it is usual to employ the beam-switching technique (cf. § 5·3). This method is employed in the case of the Army G.L. equipments, and elevation-finding with an accuracy of a few minutes of arc is possible.

For most other purposes it is convenient to scan the beam in elevation continuously, and then a special display is necessary. The most common is a range-elevation scan with brightness modula-

tion. With this lines of constant height are hyperbolae. A more realistic display is obtained by applying the range trace to the X plates (or X coils) of the c.r.t. in the normal way, and at the same time injecting these signals into the rotor of a generator geared to the elevating beam, the output from the stator winding being applied to the Y plate (or Y coils). The net result of this is that the coordinates of the display are now $r\alpha$ and r (fig. 6·5) and lines of constant height are approximately straight.

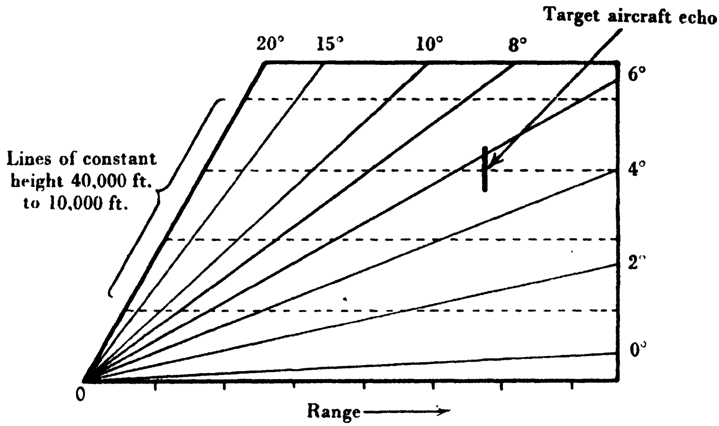


Fig. 6·5. Range-height display for V.E.B. systems.

The V.E.B. method of elevation-finding gives a satisfactory performance only when the elevation is greater than half the vertical beam width. For lower angles of elevation ground reflexions complicate the performance and the results are inaccurate.

6·5. Special methods

An interesting variation of the amplitude-comparison method consists in using two receiving aerials with different 'free-space' (or intrinsic) vertical polar diagrams* [say $F(\alpha)$ and $f(\alpha)$] at the same height h above ground. Using the same notation as before

$$E_1 = F(\alpha)\sin(hk\sin\alpha) \quad \text{and} \quad E_2 = f(\alpha)\sin(hk\sin\alpha),$$

so that the height-finding ratio in this case is simply $F(\alpha)/f(\alpha)$, i.e. the ratio of the two intrinsic polar diagrams. Methods of this

* The arrays must be symmetrical.

type have two main advantages: (i) they can be used at very low angles of elevation, and (ii) the performance is not so markedly dependent on the site.

In practice only one aerial is necessary, and this may be used both for reception and transmission. Fig. 6·6 illustrates a system consisting of a parabolic reflector, and three feed points, one (*A*) at the

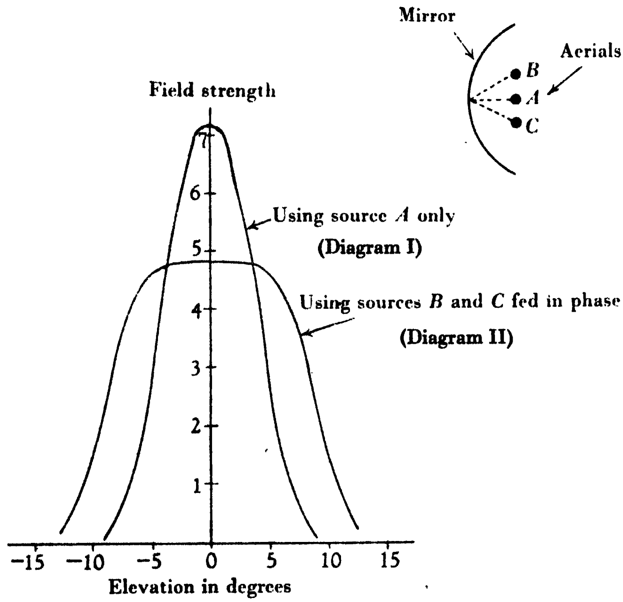


Fig. 6·6. System for height-finding at low angles of elevation.

focus, and the other two (*B* and *C*) a little above and below the focus respectively. If *A* only is fed, diagram I is obtained. If *B* and *C* are fed in parallel, diagram II is obtained. The method works from very low elevations up to an elevation given by the half-angle of the narrower beam, the lowest practicable elevation being about a fifth of the wider. It is thus complementary to the V.E.B. method.

Another interesting method of height-finding is the so-called V-Beam System. Two transmitters and two aerials are used. The first aerial produces a beam narrow in azimuth with a cosecant elevation law in the normal way. The second aerial produces a similar beam but in a plane inclined at 45° to the vertical, the planes of the two beams intersecting in a horizontal line. The two

aerials (and this line) rotate together at a uniform speed of about 6 r.p.m. in such a way that the vertical fan beam comes first. It is quite clear that a very low target would be in the intersection of the two beams and would appear as a single echo. A high aircraft, however, would be seen in the vertical beam first, and there would be a delay of a fraction of a revolution before the target appeared in

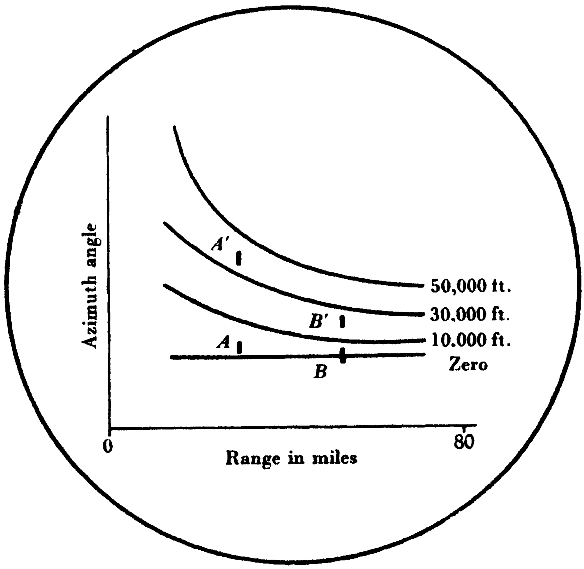


Fig. 6·7. Figure shows display with V-beam system. In the case of a low aircraft, the two responses are close together (see *B* and *B'*), whereas a high-flying aircraft gives rise to more widely separate responses (*A* and *A'*). In reading a height the height scale, which is movable, should be set so that the zero height bisects the lower echo.

the 45° beam. By making a suitable display of the indications which appear in the two beams it is possible very simply to determine the height of any aircraft. Since the rotation is continuous, height can be found continuously in much the same manner as the plan position information is supplied. The method of display is illustrated schematically in fig. 6·7, and the aerials are shown in Plate V.

Chapter 7

SYSTEMS FOR DETERMINING BOTH AZIMUTH AND ELEVATION— TWO-DIMENSIONAL SCANNING

For many radar purposes it would be desirable to be able to determine all three coordinates of the target with one equipment, i.e. azimuth and elevation as well as range. Flood-lighting type stations can do this, but can only measure one angular coordinate of one target at any one time. Stations like G.C.I., which are beamed in one plane, may be enabled to determine the second angular coordinate, but the method has various practical limitations (cf. § 6·3). The obvious method is to use a pencil-beam radar scanning in two dimensions, particularly since by using narrow beams long range and high accuracy in angle should be attainable. However, this method has its fundamental limitations, as will be seen in the present chapter.

It should be realized that the permissible scanning speed is limited, even for a station scanning in one dimension only, such as a reporting radar having a fan beam rotating in azimuth. First, it is essential that the aerial should remain pointing at the target long enough (i.e. for a period $> 2r_m/c$) for the radiated pulse to return from the target. Secondly, it is generally necessary, and this is the more stringent condition, for several pulses to be received during the period the aerial is looking at the target. This is necessary not only to ensure that the display will be satisfactory on weak signals, but also to give angular accuracy—it is generally considered that five pulses within the beam width is the minimum, and this gives an angular accuracy of about one-fifth of a beam width. For an equipment having a beam width* of θ_b and a pulse repetition rate of f_p the scanning speed $\dot{\theta}$ must not exceed $0{\cdot}2 f_p \theta_b$, which for $\theta_b = 4°$ and $f_p = 400$ sec.$^{-1}$ gives 320° per sec., or 53·3 r.p.m. This limitation is of no consequence, since scanning speeds above 10 r.p.m. (i.e. picture repetition every 6 sec.) are not required for search radars. For two-dimensional scanning stations, however, the limitations are of real importance.

* Within (say) the 6 db. down contour.

7·1. General considerations—two-dimensional scanning

In the two-dimensional scanning case, the above condition applies to the scanning rate in the coordinate θ parallel to the scan direction. For an accurate determination of the perpendicular coordinate, α, say, it is necessary to arrange to have a scan about every $0{\cdot}2\alpha_b$, where α_b is the beam width in the α direction. We see therefore that we require some 25 pulses within the solid angle, ω say, subtended by the beam, so that if a total solid angle Ω is required to be scanned, a total number of at least $25\Omega/\omega$, or a time of at least $25\Omega/\omega f_p$, must elapse before a complete picture can be obtained. In actual practice most scans are inefficient in the sense that some time may be lost in scanning outside the required area, or in scanning some parts of this area unnecessarily densely. We may therefore write a factor k such that at least $k\Omega/\omega$ pulses are required for each scan, k having a value depending on the type of scan used, but being never less than 25. If the coordinate α does not need to be determined to the highest possible accuracy it is possible to relax this condition somewhat; we may, for example, increase the angle between scans to $\frac{1}{2}\alpha_b$, in which case k for a perfectly efficient scan drops to 10. It must be remembered that in scanning any appreciable solid angle it is most difficult to design an efficient scan, keeping the lines of scan a constant angular distance apart, for the same reasons which make it impossible to have a perfect map projection on a flat surface. Zigzag scans are also necessarily inefficient; therefore unless we have to scan the whole 2π of one coordinate, mechanical scanning arrangements should be avoided whenever possible—they must either zigzag or waste time on unwanted values of the angle concerned.

We see then that the maximum picture repetition rate (N, say) cannot exceed a certain value

$$N \leqslant \frac{f_p \omega}{k\Omega}, \tag{7·1}$$

while f_p must be low enough to allow an echo from the greatest range r_m to appear on the time base, so that, allowing say 20 % for fly-back time,

$$\begin{aligned} f_p &\leqslant 0{\cdot}8c/2r_m \\ &\leqslant 0{\cdot}4c/r_m. \end{aligned} \tag{7·2}$$

Combining these we obtain

$$N \leqslant \frac{0{\cdot}4c\omega}{kr_m\Omega}. \tag{7·3}$$

In these formulae we see that k is fixed by the type of scan (efficiency factor) and accuracy required, and the latter factor also fixes an upper limit to the beam size ω, so that once the area to be scanned is fixed the maximum picture repetition rate is definitely determined.

In some applications the picture repetition rate must have a certain (minimum) value. If this is so, we may determine the maximum range as follows. If the aerial puts all its energy into the beam of solid angle ω, its gain will be $G = 4\pi/\omega$, and to allow for side lobes we may write $G \leqslant 4\pi/\omega$. Combining this with equation (7·3) above, we have

$$G \leqslant \frac{4\pi \times 0{\cdot}4c}{Nkr_m\Omega},$$

and if we insert this value for G_s and G_r into equation (3·11), we obtain

$$P_m \geqslant \frac{P_s \times 0{\cdot}16c^2\lambda^2 A}{4\pi N^2 k^2 \Omega^2 r_m^6},$$

whence
$$r_m \leqslant \sqrt[6]{\left[\frac{P_s 0{\cdot}4A}{P_m \pi}\left(\frac{\lambda c}{Nk\Omega}\right)^2\right]}, \tag{7·4}$$

so that r_m is completely determined. It should, however, be remarked that P_m depends on the perception factor Π, which depends on the number of pulses superimposed at the display,* so that if k is made too low a larger received signal will be necessary than would normally be required. In fact, on a brightness-modulation display, the spot must be brighter to be detectable if it only occurs once than if the trace is repeated several times, while on a deflexion display the effect of repeated traces is to blur the noise so that again a smaller

* This effect is sometimes expressed by defining a scanning loss factor (f_s, say), such that the perception factor Π_s when scanning at any given velocity is related to the perception factor Π for slow scanning by the relation $\Pi_s = f_s \Pi$, so that $f_s > 1$. The factor is only appreciably greater than unity when the scanning speed is such that less than five or six successive traces are superimposed at the display. The relation between f_s and the scan rate, display spot size, beam width, etc., is complex (cf. Watson, D. S., *J. Instn Elect. Engrs*, vol. 93, part III A, p. 143 (1946), § 5).

As an example, the relation for a particular A.I. system is found to be

$$f_s = 1 + 40\Big/\left(\frac{\omega f_p}{\Omega N}\right)^{1{\cdot}3}.$$

Knowing the value of f_s, the maximum range can be calculated using equation (3·12). Thus if f_s is equal to 2 the maximum range is reduced by about 20% when scanning.

signal can be seen. However, once several traces are superimposed, an increase in the number has relatively little effect.

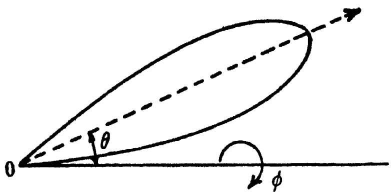
It is seen, therefore, that the two-dimensional scan method has a fundamental limitation dependent ultimately on the velocity of light. If a predetermined accuracy (and therefore beam width) is required, the picture repetition rate may become so low as to be a serious operational disadvantage, while if the picture frequency is kept up the maximum range and the angular accuracy both have to be reduced. How this works out in practice will be considered in §§ 7·3–7·4 below.

7·2. Practical scanning systems

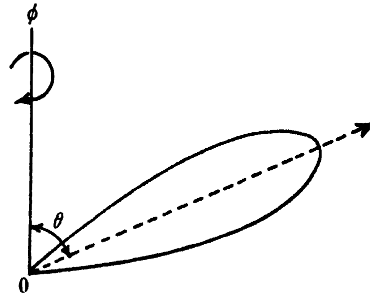
Several different scanning systems are employed in radar. The most common are: (i) conical scanning, (ii) helical scanning, (iii) conical spiral scanning, and (iv) precessing conical scanning. These methods are illustrated in fig. 7·1. The first, though it scans in two angular dimensions, is really only a one-variable scan, and has effectively already been considered as a beam-switching method (cf. § 5·3). It will be further considered in § 7·2·2 below. Helical scanning may be done in either of two ways, either a rapid scan in azimuth with a slower elevation wobble, or vice versa, the former being the more usual. It corresponds to the normal television scan, although it may extend over a full 360° of azimuth. The precessing conical scan, the precession being in azimuth, is an alternative to the helical scan, while the conical spiral scan is a system with complete symmetry about the mean line of shoot.

Considering first the last-mentioned (conical spiral), this method is of interest since it was used in the first centimetric A.I. equipment to be produced. The beam is produced by means of a paraboloidal reflector whose axis rotates at 1000 r.p.m. about the line of flight of the aircraft and simultaneously oscillates at a frequency of about 1 per sec. between the angular limits of $1\frac{1}{2}$ and $31\frac{1}{2}°$ relative to the line of flight. The dipole and reflector feeding this mirror are stationary, so that the beam starts describing a circle of 3° radius and spirals out until it is describing a 45° radius circle and then spirals in again. The beam width is about 6°, the diameter of the mirror aperture being 2 ft. and the wave-length 10 cm.

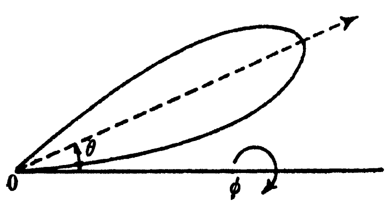
Helical scanning has also been used for A.I. equipments, the scanning lines being horizontal. A 180° scan in azimuth has been



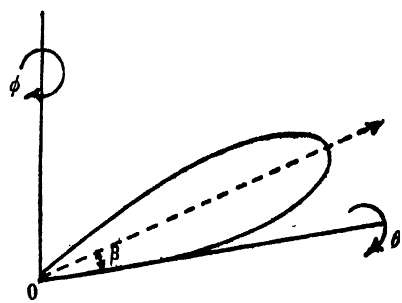
(*a*) Conical scanning. $\theta=$constant, $\dot{\phi}=$constant



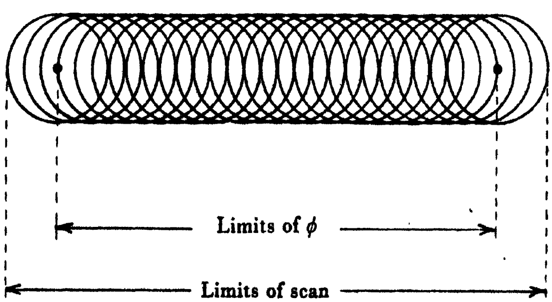
(*b*) Helical scanning. $\dot{\phi}=$constant (fast scan), θ varies slowly between predetermined limits (limits usually include $\frac{1}{2}\pi$)



(*c*) Conical spiral scanning. $\dot{\phi}=$constant (fast scan), θ varies slowly between predetermined limits (one limit ~ 0)



(*d*) Precessing conical scanning. $\beta=$constant, $\dot{\theta}=$constant (fast scan), ϕ varies slowly between predetermined limits (cf. fig. (*e*))



(*e*) Illustrating precessing conical scanning

Fig. 7·1.

used with two mirrors back-to-back so that the efficiency of the scan is maintained. The whole system is then tilted cyclically between the required elevation limits. A similar method has been proposed for a ground station, the lower limit of elevation in this case being a fraction of a beam width above horizontal. The precessing conical scan may be used as a substitute for the helical scan when only a limited azimuth coverage is required. The semi-vertical angle (β) of the cone is made equal to half the angular range of elevation required. This system possesses the advantage over a helical scan with the scan lines vertical (or a zigzag scan with the lines nearly vertical) that the scan is extended by 2β in azimuth beyond the mechanical limits of scan (cf. fig. 7·1 *e*).

When 360° in azimuth and a limited range in elevation has to be covered the helical scan is the most efficient. For other cases the choice will vary according to the operational requirements, but it should be noted that for A.I., which is a short-range equipment, efficiency of scan is not of prime importance.

7·2·1. *Display systems*

The display system for the conical spiral scan is of interest. It is shown in fig. 7·2, being a P.P.I. type of display in r and ϕ (fig. 7·1 (*c*)) with no representation of θ. The range increases from the centre outwards, and while a target on the line of flight gives rise to a complete circular echo (i.e. a response is obtained for all values of ϕ when θ is very small), a target displaced at a small angle gives a semicircle, and as the displacement increases the length of the arc shortens since the beam scans over the target more rapidly when θ is large. The direction of the angular displacement (left, right, up, or down) is indicated directly on the display by the direction of the centre of the arc on the cathode-ray tube.

For the helical or precessing conical scan a television (azimuth-elevation) display is of course possible, but range has to be indicated on a separate presentation unit. The auto-ranging device mentioned in §4·1·1 may have value here, so long as only one target is in view (as is normally the case with A.I.); it is then arranged that the television scan is 'gated' in range so that only responses in a small interval corresponding to the range required are shown on the display. An alternative arrangement is to have separate range-azimuth and range-elevation displays, and often the third

coordinate not shown in any display is 'gated' by manual adjustment from the other display. Most of these equipments need a separate operator in the aircraft in addition to the pilot, which may be a serious disadvantage.

One further factor which needs to be considered when deciding on the display is the scale of the picture. If the scale is too large echoes will be presented not as continuous lines but as a chain of

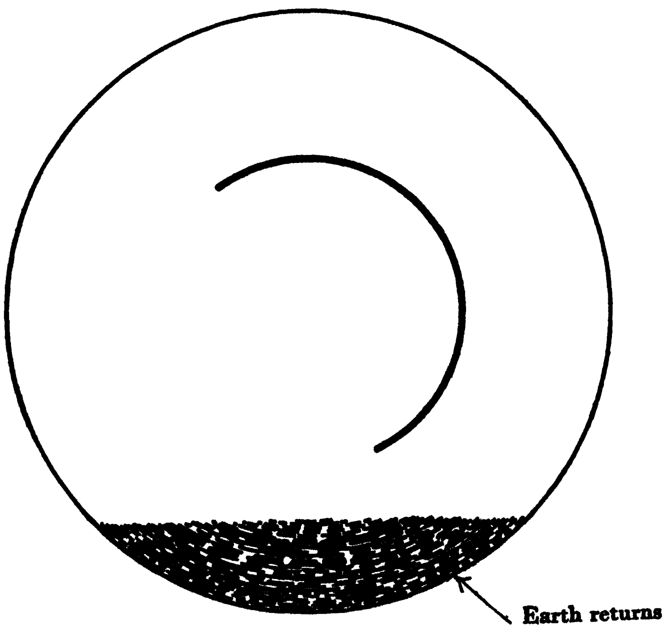


Fig. 7·2. A.I. with spiral scan display. The echo response as shown indicates that the target aircraft is starboard and 'above'.

dots. The spot size with the normal cathode-ray tube is about 1 mm., so that if the display is of the range-azimuth type the scale in azimuth should not be greater than $10\theta_b/n$ radians per cm., where θ_b is the beam width and n is the number of pulses per beam width.

For reporting radars (e.g. ground or shipborne) the problem of displaying on a two-dimensional screen the three coordinates of a large number of echoes presents considerable difficulties, and a P.P.I. display together with a range-elevation display are favourites.

Such a system has not been greatly used, however, on account of the difficulties indicated in §7·3 below. The displays for such a system have therefore not been fully developed.

7·2·2. *Auto-follow systems*

For auto-following (cf. §4·1·1) the simple conical scan is used, as explained in the section on beam-switching techniques (cf. §5·3). The axis of the system is caused to aim at the direction of the target by applying the misalinement voltages derived from the 'split' signals to a suitable servo-system operating the elevation and bearing drive motors.

In one system which is widely used the drive motors are composite machines each comprising a split-field motor and a tachometric generator. The split-field motor carries a constant current in its armature, and the direction of rotation and torque are controlled by currents in the field windings. These currents are derived from the misalinement voltages after suitable amplification. The output of the tachometric generator is a voltage proportional to the velocity of rotation and is fed back into the input of the amplifier with the misalinement voltages. One special feature of this velocity feed-back system is that if the signals fade out the system still remains stable and continues to move with a constant velocity.

7·3. Long-range radars using two-dimensional scanning

As an example let us consider the design of a long-range search radar. We will suppose that an r_m of 150 miles is required with a picture repetition frequency of 10 per min. and a solid angle of search of π (i.e. a quadrant of a sphere, or 180° in azimuth with 0–90° in elevation).

From equation (7·2) we see that f_p cannot exceed 500 per sec. for $r_m = 150$ miles, therefore from equation (7·1) with $\Omega = \pi$, ω must not exceed about 0·01 steradian, even if we take k as low as 10. Pencil beams more than about 6° wide, or gains exceeding 1000–1500, are therefore excluded. It is usually necessary to obtain considerably greater precision than a 6° beam allows, and generally an aerial gain in excess of 1500 is required to give a maximum range of 150 miles, remembering that reflexions from the ground cannot be used to increase the range. In fact, if we put $P_s = 10^6$ W., $P_m = 10^{-12}$ W. (cf. §2·6), $\lambda = 10$ cm., $A = 10$ m². in equation (7·4), together with the values of N, k and Ω given above, we obtain

$$r_m = 186K_m = 116 \text{ miles.}$$

Since $P_s \propto r_m^6$ we see that to attain a maximum range of 150 miles a sender power of 4·7 MW. would be necessary, and even this is with a minimum of interlacing and a perfectly efficient scan ($k = 10$). For $k = 40$ (allowing some inefficiency on a full-accuracy scan) r_m is reduced to 72 miles (for a 1 MW. sender). It must be concluded, therefore, that with powers at present available a long-range search radar with two-dimensional scanning is impracticable unless an upper range limit of some 70 miles is acceptable.

In practice we find that the wartime long-range search radars were either flood-lighting stations or stations beamed in azimuth only, with only a very few exceptions. For some purposes (cf. § 8·2) the increased resolution which the pencil-beam radars provide is of first importance, and short-range radars of this type, providing for search with a maximum range of say 70 miles, have been designed. A common form is that using rapid elevation scan (cf. § 6·4), with a simultaneous rotation of the whole system about a vertical axis.

7·4. Air-interception radars

As an example of two-dimensional scanning it may be useful to consider the design of an A.I. radar, an equipment which was installed in fighter aircraft during the war as an aid to the interception of other aircraft. A.I. is a relatively short-range device providing information about the range, bearing and elevation of the target, and often it is sufficient to use a beam which scans the region of the sky ahead of the aircraft only. The analysis given in § 7·1 applies, but there is an additional limitation to the range due to ground returns. For example, when the beam is looking horizontally ground returns will begin at range equal approximately to $60h/\alpha_b$, where α_b° is the vertical beam width, and h the height of the aircraft. For $\alpha_b = 15^\circ$, for example, and an aircraft height of 1000 ft., the range limitation will be slightly under 1 mile, while at 10,000 ft. it will be about 8 miles. This limit may only be increased by using a narrower beam; it also varies with the elevation of the mirror, being of course worse during the part of the scan when the beam is tilted downwards.

Supposing we wish to design an A.I. having a maximum range of 8 miles, it is not sufficient to choose f_p according to equation (7·2), since ground clutter may be received from much greater ranges and would appear on subsequent time-base sweeps. It may be, for

instance, that a range of 8 miles on a target aircraft corresponds to a range of 100 miles on ground targets. Accepting this, a pulse repetition rate of 800 c./sec.* will be needed to avoid these unwanted echoes from distant targets complicating the 'following' of aircraft at shorter ranges.

We will assume a picture repetition rate of 30 per min. is required, and that $\alpha_b = 10^\circ = \theta_b$, so that $\omega = 0{\cdot}025$ approx. If we use widely spaced scans ($k = 10$), then equation (7·1) gives $\Omega \leqslant 4$ for $f_p = 800$, which means that we can scan just under a hemisphere ($\Omega = 2\pi$). This would be barely satisfactory, leaving no margin for inefficiency of the scanning system, so that we see that 10–12° beams are about as small as are allowable. It remains to be seen whether the scanning rates are practicable, and whether sufficient range can be achieved.

The scanning speed in the 'fast' direction is given by one beam width (say 10°) per five pulses, i.e. 2° per pulse, or 1600° per sec. This is 267 r.p.m., which is quite feasible. For maximum range we use equation (3·12), with $\lambda = 10$ cm., $G_s = G_r = 4\pi/w = 500$ approx., $P_m = 10^{-12}$ W., $r_m = 8$ miles $= 12{,}850$ m., $A = 10$ m.^{2}, and obtain $P_s = 2{\cdot}15$ kW. as the required sender power. In practice rather more than this will be necessary, to allow for feeder losses, poor perception factor on an open scan, etc., but it is clear that adequate range is attainable without difficulty, and this would still be true if a smaller mirror giving a 15° beam were used. Such a wider beam would, of course, reduce the accuracy in angle but would make the scanning problem easier—with the 10° beam the most that could be satisfactorily scanned with the required picture frequency would be a cone of semi-vertical angle 45° ahead of the aircraft (corresponding to $\Omega = 1{\cdot}84$ steradian, allowing a k of about 20 for the scan).

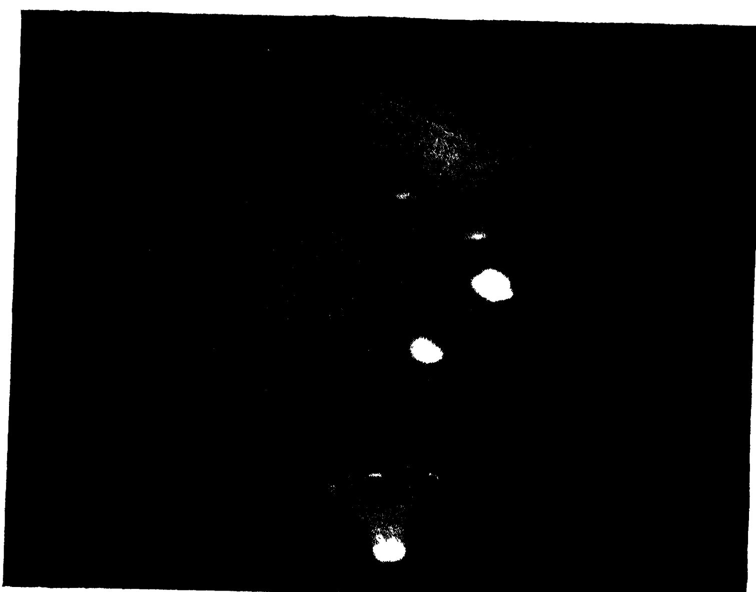
* Higher repetition rates than this have been used, even with A.I. radars capable of 'seeing' ground targets at ranges of the order of 100 miles, and the presentation has then been complicated by the presence of ground returns appearing on the second and third trace. These unwanted echoes can be resolved by momentarily jittering the pulse repetition frequency, and it is possible that a continuously jittered pulse repetition frequency might randomize distant echoes sufficiently to allow f_p to be increased above the value given by equation (7·2), but this technique still remains to be developed.

Chapter 8

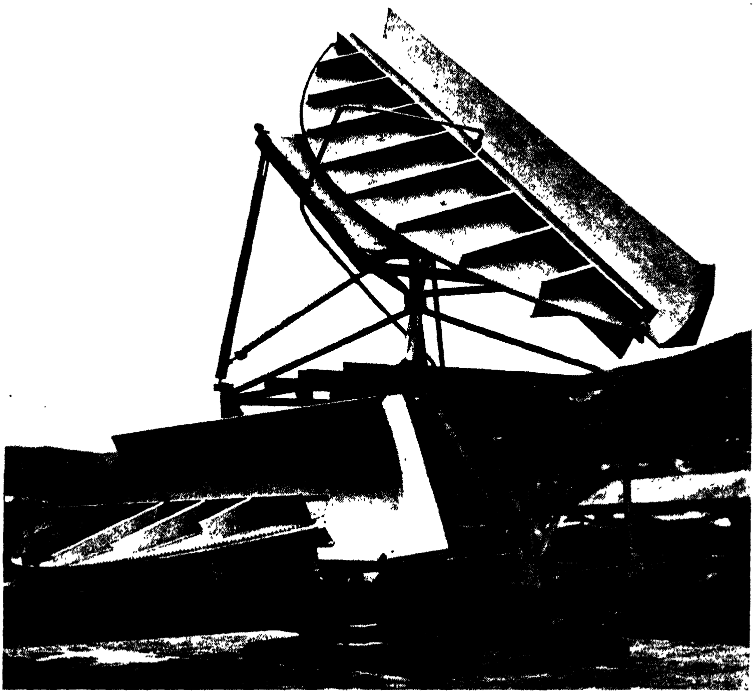
UNWANTED ECHOES

We have already seen that the radar detection of ships on the sea, tanks on the land and even aircraft in the sky is complicated by the simultaneous observation of echo signals scattered from the surface of the sea in the first case, from land irregularities in the second, and from one or both of these in the third. In this chapter we discuss the origin of these unwanted echoes and consider how their effect can be reduced to a minimum.

An interesting example is the 'long-range scatter' obtained with the 12 m. C.H. stations. These stations operate on a wave-length in excess of the critical wave-length for piercing the E-layer at oblique incidence, and consequently, under suitable conditions, strong reflexions due to waves which have travelled to and from the radar station via the upper atmosphere are obtained from points on the earth's surface beyond the 'skip distance'. These reflexions are from considerable ranges, and so it is only necessary to use a very low repetition frequency to ensure that they cause no interference with the tracking of aircraft at the shorter ranges. Thus if the pulse repetition frequency is 25 c./sec., the time between consecutive pulses is 40 msec. In order to track aircraft up to a maximum range of 200 miles, it is necessary to arrange that the cathode-ray tube spot completes its journey across the face of the display tube in 2 msec., so that the black-out time is 38 msec. The scatter usually originates from points on the earth's surface at ranges from the station between 2000 and 4000 miles, and so it arrives during the black-out period and is not seen on the following trace. Under phenomenal conditions when the scatter is returned from ranges in excess of this, the pulse repetition frequency is reduced to 12·5 c./sec. Fortunately, long-range scatter of this type is not obtained with the higher radio-frequencies, but, on very short wave-lengths (1 m. and less), reflexions are obtained from cloud formations. Very strong reflexions are, in fact, obtained from cumulo-nimbus clouds, and the discovery of this phenomenon has resulted in the use of airborne radar equipments to aid pilots in navigating around these dangerous



Photograph showing clouds as viewed from aircraft (upper plate) with the corresponding indications as observed with a 3 cm. airborne radar (lower plate).



The V-beam radar aerial system (cf. § 6·5). 'Cheese' aerials are used to feed power into cylindrical reflectors designed to produce a cosecant law (cf. § 5·4) in the plane perpendicular to the plane of the 'cheese'. The lower aerial produces a vertical 'fan' beam (giving accurate azimuth determination), while the upper aerial gives an inclined beam. From the azimuth at which an echo is obtained with the latter aerial, the angle of elevation of the target can be deduced, as described in the text.

It is interesting to realize that the upper part of the reflector profile is approximately parabolic, concentrating an appreciable part of the total radiated energy into a plane wave front, which by diffraction yields a normal-shaped beam almost horizontally (like the full-line beam of fig. 5·1). The lower part of the reflector directs the remaining energy upwards to produce an over-all diagram more like fig. 5·11.

Note. 'Cheese' aerials resemble wave guides in that the radiation is confined between parallel conducting sheets, but have a parabolic reflecting surface in the other plane giving rise to a plane wave front at the mouth when fed from a 'point' source such as the mouth of a wave guide.

storm centres. It has also resulted in the use of ground radar equipments for supplementing meteorological information obtained in other ways. When, however, our purpose is the location of ships or aircraft, reflexions from clouds interfere with the operation of the radar and, in consequence, much thought has been given to the problem of minimizing this effect. One of the most interesting methods devised for the purpose is that using crossed polarization. The principle of the method is as follows. Separate aerials are used for sending and receiving and the polarizations are crossed (i.e. if the sender is arranged to transmit a vertically polarized wave, the receiver is arranged to be sensitive only to horizontally polarized waves). With such a system weak, but nevertheless detectable, signals are obtained from targets of large and complex shapes such as ships and aircraft, but practically no signal is obtained from cloud formations, which behave as a conglomeration of small scattering particles.

8·1. Sea clutter

When short radio waves strike the surface of the sea at small angles of incidence, most of the energy is specularly reflected unless the sea is very rough (cf. § 3·1·2). A small amount of energy is, however, scattered in all directions. If the sea is illuminated by a radar set sufficient energy may be returned to produce echoes extending to a distance of several miles. The range up to which the sea echoes are observed depends, of course, on the performance index of the radar and on the state of the sea. It also depends on the polarization used. For wave-lengths of 10 cm. or more the sea clutter is much stronger with vertical polarization than with horizontal polarization; for wave-lengths of 3 cm. or less the difference is small. It was largely for this reason that vertical polarization was not adopted for coast-watching stations operating on the longer wave bands.

Sea clutter can be a fundamental limiting factor in the detection of small targets. It can also prevent the detection of echoes which are stronger than the sea clutter where the display is of the P.P.I. type and all the signals are limited at just above noise. To overcome this, a number of special circuits have been devised which reduce the receiver gain by the appropriate amount at each range so that the amplitude of the clutter is less than the saturation level. In one of these, the receiver gain is adjusted according to an assumed law for the variation of clutter with range. With a rotating-beam radar

system, particularly when high rotational speeds are employed, the method is unsatisfactory since it cannot compensate for variations in the strength of the clutter with the direction of the aerial relative to the wind. A method which is satisfactory in these circumstances employs a quick-acting automatic gain control responding to the clutter but not to a single stronger echo. This device has been used on a number of naval sets,* for which a response time of about 4 times the sender pulse duration and a decay time of about 25 μsec. was found to give the best performance. The A.G.C. method has the advantage that it is automatic and compensates for variations in the strength of the clutter as the aerial rotates. Nevertheless, both methods are palliatives only and cannot be regarded as providing anything like a complete cure.

8·2. Ground clutter

Radar responses are also obtained from grounded objects, and in some cases, e.g. in the case of the G.C.I. stations, ground clutter limits the performance. For this reason G.C.I. stations are usually sited very carefully, the preferred location being at the centre of a 'saucer'-shaped site, of about 5 miles diameter. In these circumstances the ground clutter is confined to relatively short-range responses, and the tracking of aircraft at longer ranges is not complicated by the presence of unwanted echoes. It should be noted, however, that good sites are not always available, and when ground radars are required in mountainous country special methods have to be employed. The devices mentioned in §8·1 are of value here, but their application is somewhat limited. A more satisfactory answer is to use a very narrow-beamed radar† with a magnified display. Further improvement is possible using a Flicker display. This consists in using effectively two displays, one of which presents the picture as seen on one rotation of the aerial and the other on the next rotation. A viewing system is provided so that these two pictures are seen alternately, superposed and in rapid succession, with the result that aircraft responses which do not coincide exactly from one picture to the next appear to 'flicker', whereas responses from grounded objects do, and therefore they fade into the general background. Several methods have been used for producing flicker

* Ross, A. W., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 236 (1946).
† Cf. §7·3.

displays of this type, including one in which the display is photographed every rotation, the film being rapidly developed and fixed and, whilst still wet, the two consecutive pictures are viewed simultaneously, the picture moving on at each rotation of the aerial.

Perhaps the most valuable method of minimizing the effects of ground (and sea) clutter depends on the application of Doppler's principle. This is discussed in the following section.

8·3. Doppler systems

It is well known that the frequency of waves reflected from a moving target exhibit the Doppler frequency shift. The observed frequency f' from an approaching target is given by

$$f' = f\left(\frac{c+v}{c-v}\right), \tag{8·1}$$

where f is the radiated frequency, c is the velocity of the waves, and v is the relative velocity of the target. The Doppler frequency shift is therefore given by

$$\Delta f = \frac{2v}{c-v} f = \frac{2v}{c} \times f \text{ approximately.} \tag{8·2}$$

It follows that the Doppler frequency Δf is proportional to v, the radial component of the velocity of the target relative to the radar (or in the case where a ground radar is used, to the radial velocity of the target), and to f, the frequency used by the radar system.

The simpler application of Doppler's principle is to stations which radiate a continuous wave of constant amplitude and frequency. Such systems usually have separate sending and receiving aerials. The waves reflected from the target arrive at the receiving system together with some stray radiation from the sender. When the target is moving the reflected wave differs in frequency from the stray radiation by an amount equal to the Doppler frequency and the two received signals interact producing 'beats', the resultant signal being effectively amplitude-modulated at the Doppler frequency. The receiver is similar to a conventional broadcast receiver,* containing a detector and audio-frequency stages, the

* If the maximum radial speed of the target is 400 m.p.h., the maximum Doppler frequency is 3600 c./sec. for a set operating on the 3000 Mc./sec. band. The effective band width of the receiver must therefore be about 4 kc./sec. as compared with 4000 kc./sec. for a pulse system. Consequently, to attain comparable sensitivity with a c.w. system the radiated power need only be $1/10^3$ of the peak power of a pulse system with the same aerial gain.

output being applied to a pair of telephones. If the Doppler frequency is in the audible range it is heard as a continuous note in the telephones. This simple application provides no measurement of range other than is possible from an estimate of the amplitude of the received signal, and is therefore not a radar system according to our definition. Nevertheless, such systems have proved of value for warning purposes.

In applying the Doppler principle to pulse radar systems it is generally desirable to make the successive pulses 'coherent', i.e. phase-related to a steady c.w. signal. Methods in which this is not essential have been used, but their application is restricted to cases where there is some 'clutter' due to stationary objects at the same range as the target, with which the Doppler beat can be observed. These systems have other disadvantages and will not be described further.

In normal radar senders, the relative phases of the r.f. wave in successive pulses vary in a random manner because the oscillations which build up in the sender when the modulating pulse is applied are excited by random voltages arising from thermal agitation. The relative phases of successive pulses can, however, be made coherent by feeding a small c.w. voltage into the oscillator of the sender since the oscillations in the pulse will then build up from this c.w. voltage. In consequence, all the pulses start in the same phase relative to the c.w. If now the pulse reflected from a target is mixed with some of this c.w. in the receiver, the phase difference between the two will depend on the range of the target. Therefore, the echo from a fixed target will remain steady in amplitude at the output of the receiver, but the echo from a moving target (e.g. an aircraft) will beat owing to the Doppler effect.

In practice it is convenient to use a system which presents only the responses from moving targets. This is achieved by delaying the video signals by a time equal to the repetition period, and subtracting the delayed from the undelayed video signals. A supersonic cell* is used to give the delay. The output of such a system depends on the difference in amplitude of the video signals in adjacent recurrence periods. On the assumption that the changes in amplitude are entirely due to the radial velocity of the target, it is clear that fixed echoes will not appear at the output. Echoes from moving

* Cf. § 4·2·2.

targets will, however, show a sinusoidal beat, the peak amplitude being dependent on the radial velocity. The peak amplitude of the cancelled picture will be a maximum for a target moving an odd number of quarter wave-lengths per recurrence period. The corresponding velocity in m.p.h. is given by $f_p \times 170n/f_0$, where f_p is the pulse repetition frequency in c./sec., f_0 is the radio frequency in Mc./sec., and n is an odd integer. At this velocity, the subtraction will give a double amplitude pulse, whereas the noise voltage will be increased by $\sqrt{2}$; thus a peak signal/noise improvement of $\sqrt{2}$ is

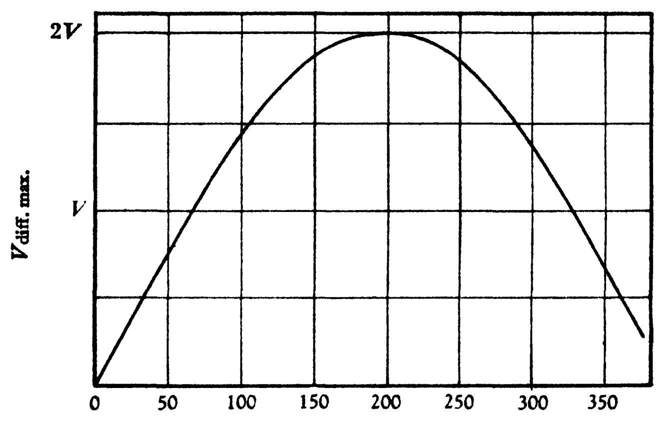


Fig. 8·1. Velocity diagram for $\lambda = 60$ cm., $f_p = 600$ c./sec.

obtained. It is apparent that in practice one must adjust the repetition frequency according to the range of radial velocities over which a useful performance is required. This is illustrated in fig. 8·1, which shows the velocity diagram corresponding to a sender operating on a wave-length of 60 cm. ($f_0 = 500$ Mc./sec.) and a pulse repetition frequency (f_p) of 600 c./sec. It will be noted that a peak signal/noise improvement of about $\sqrt{2}:1$ is obtained for radial velocities of about 200 m.p.h., but for radial velocities less than 50 m.p.h. or between 350 and 400 m.p.h. the received signal strengths are very poor.

To obtain good cancellation the following conditions are necessary: (i) a very high degree of sender frequency stability, (ii) accurate phase-locking to the reference c.w. oscillator, and

(iii) a very high degree of frequency stability in the reference c.w. oscillator. In practice the cancellation obtained is limited by certain other factors. If a beamed rotating aerial system is used, it is easy to see that consecutive pulses received back from a target at fixed range will not in general be equal. This is due to the movement of the aerial beam across the target, and it may be shown that if the electronic equipment is perfect, cancellation better than about $1200N/\theta f_p$ % cannot be achieved. In this formula, N is the rotational speed of the aerial in rev./min. and θ is the beam width in degrees. As an example, consider a ground radar equipment with an azimuthal beam width of 5°, aerial rotational speed of 6 rev./min. and a recurrence frequency of 600 c./sec. The limitation is $1200 \times \frac{6}{3000} = 2\frac{1}{2}$ % approx. Nevertheless, a reduction of signal strength of this order is very valuable and is usually good enough for all but the worst sites.

Chapter 9

ELECTRICAL CHARACTERISTICS OF TYPICAL RADAR EQUIPMENTS

In this chapter we give the design constants of certain radar equipments which have been used during the war, mainly to illustrate the types of design which are possible for various purposes. Other factors, such as the choice of wave-length (cf. § 3·5), are also illustrated. Although a few examples of shipborne radar have been included, it must be emphasized that they form a specialist field, where design is considerably influenced by such factors as the height of the masts available and the permissible loadings.

9·1. Ground radar

(1) C.H. ('Chain Home') stations were the first search radars to be built and used operationally. Using wave-lengths of about 12 m., they employ the flood-lighting principle (cf. § 5·1), the aerials being mounted on towers 70–110 m. high. Azimuth is found with crossed receiving dipoles and a goniometer (cf. § 5·1·1), and elevation by a similar method using spaced aerials (cf. § 6·2·1), both being null settings on a range-amplitude display (cf. § 4·1). Azimuth and elevation are both markedly dependent on site (flight calibrations being usually essential). On poor sites elevation measurement is impossible (fairly level ground for about a mile in front of the station being needed). Two height systems are often provided, 'A' (aerials at 24 and 72 m.) for elevations $1\frac{1}{2}$–6°, and 'B' (24 and 12 m.) for 6–15°. Azimuth accuracy is about $\pm 2°$, and elevation about $\pm\frac{1}{4}°$. With $P_s = 4 \times 10^5$ W., $P_m = 2{\cdot}5 \times 10^{-13}$ W. (cf. example, p. 51, but using $B = 150$ kc./sec.) due mainly to aerial noise, $G_s = 20$ (a vertical stack of dipoles with aperiodic reflectors), $G_r = 3{\cdot}3$ (two crossed dipole pairs, one $\frac{1}{2}\lambda$ above the other and connected in phase), $f_p = 25$ c./sec., $\tau = 5$–$20\,\mu$sec., $B = 50$–500 kc./sec. (three settings), we obtain $(r_m)_{\text{f.s.}} = 100$ km.* On a typical site ranges might be 90 miles on an aircraft flying at 15,000 ft., but only 50 miles

* All ranges are for an average medium aircraft ($A = 13$ m.²).

at 5000 ft. The poor performance on low-flying aircraft is the principal defect of these stations.

(2) G.L. ('Gun-laying') stations operate on a wave-length in the band 3–5½ m. These stations use a modified Adcock azimuth system (cf. § 5·1·2), giving an accuracy of about $\pm \frac{1}{2}°$. Two dipoles at λ and $1{\cdot}5\lambda$ above the ground are used to determine elevation (accuracy $\pm 1°$ in the range 20–40°), using a goniometer. A range-amplitude display is provided, the 'cross-wire' method (cf. p. 60) being used to give high accuracy. Azimuth and elevation are null settings, the whole equipment rotating for azimuth determination. With $P_s = 8 \times 10^4$ W., $P_m = 10^{-13}$ W., $G_s = G_r = 3$, $f_p = 1000$ c./sec., $\tau = 1$–$3\,\mu$sec., we obtain $(r_m)_{\text{f.s.}} = 32$ km. or 35,000 yd. Ground reflexions occur but do not sensibly increase the range; near null settings, in fact, the signal is weak and the working range is more nearly 25,000 yd.

(3) C.H.L. ('Chain Home Low') stations were built to supplement the C.H. at low elevations; similar equipments were used to detect surface vessels. The wave-length used is 1·5 m. and the aerial is beamed, consisting of a 32-element broadside array used both for sending and receiving (cf. p. 74). $G_s = G_r = 80$ approx., $P_s = 10^5$ W., $P_m = 10^{-13}$ W., approx. (cf. examples, pp. 24, 50), $f_p = 400$ c./sec., $\tau = 3\,\mu$sec., giving $(r_m)_{\text{f.s.}} = 100$ km. Sited on cliffs or towers overlooking the sea, $\phi = 2$ in the lobe maxima, giving ranges of about 130 miles, but with interference gaps. Generally a P.P.I. display is employed, but beam-switching and a range-amplitude display is used when accurate azimuth (i.e. $\pm \frac{1}{4}°$ or less) is required. C.H.L. stations provide no elevation-finding.

(3*a*) G.C.I. ('Ground Control of Interception') stations used C.H.L. equipment with different aerials. The 'mobile' type had a single C.H.L. aerial with its centre 2λ above the ground, and was divided into an upper and a lower half for spaced-aerial elevation-finding (cf. § 6·3 and Plate III A). The range is slightly reduced as compared with C.H.L., (*a*) because $G_r = 40$ instead of 80, due to using only half the array, and (*b*) because the sending and receiving interference maxima do not occur at the same elevation angle so that ϕ never attains a value of 2. Some G.C.I. (the 'fixed') stations had an additional four-stack aerial at 5λ above the ground to give elevation-finding down to $1\frac{1}{2}°$ and better ranges at low elevations.

(4) The Fighter Direction station was one of the few 50 cm. ground radar stations which was produced in the war, and is of interest as it uses two-dimensional scanning (cf. §7·3). With a paraboloidal aerial of wire mesh 30 ft. in diameter a $3\frac{1}{2}°$ wide beam is made to scan rapidly in elevation from 2 to 12° while sweeping continually in azimuth. The essential characteristics of this equipment are $\lambda = 50$–60 cm.*, $G_s = G_r = 2400$ (common aerial working), $P_s = 2 \times 10^5$ W., $\tau = 2\,\mu$sec., $f_p = 400$ c./sec., $P_m = 1 \times 10^{12}$ W., whence $r_m = 208$ km. or 130 miles (note that free space conditions apply since the beam is clear of the ground). The weakness of this system is in the scanning rates; note that equation (7·2) would allow $f_p < 600$ c./sec., whereas 400 c./sec. is used. The elevation scan rate is 10° and back in 0·08 sec., giving 5·3 pulses per beam width, so that the maximum azimuth scan rate for good azimuth accuracy would be one beam width ($3\frac{1}{2}°$) in 0·2 sec. (5 vertical scan half-periods), i.e. 3 r.p.m. For good tracking a plot every 10 sec. was desirable and this could only be obtained by sweeping a quadrant of azimuth at 3 r.p.m. and speeding up the aerial in the other three quadrants so that tracking outside the preferred quadrant was poor. A P.P.I. display was used, with a V.E.B. type (cf. § 6·4) for elevation determination.

(5) A typical centimetre wave equipment is the G.L. equipment, a later equipment than example (2) above and one capable of better angular accuracy. This uses a rotating eccentric dipole (cf. p. 81) at the focus of a 4 ft. paraboloid used both for sending and receiving. Automatic following (cf. §§4·1·1 and 7·2·2) is normally employed and accuracies of the order of ± 35 yd. for range and ± 8 min. for azimuth and elevation are obtained. Typical characteristics are $\lambda = 10$ cm., $P_s = 2 \times 10^5$ W., $P_m = 1{\cdot}2 \times 10^{-12}$ W., $f_p = 1500$ c./sec., $\tau = 0{\cdot}5\,\mu$sec., $G = 1100$, giving $r_m = 6 \times 10^4$ m. or 66,000 yd. Owing to the beam switching method employed the full gain is not available, and the range for which signal strength is sufficient for automatic following is approximately half that given above. A range-azimuth type of display is used with gating (cf. §4·2) and with meter indication for azimuth and elevation.

* Operationally the ability to work on any frequency within a band is advantageous. In this case the sender, receiver and common aerial unit were rapidly tunable and the aerial was of a broad-band design.

9·2. Airborne radar

(1) The first A.I. ('Air Interception') equipments to be used operationally employed a wave-length of 1·5 m. (this was the shortest wave-length then available on which an adequate sender power could be obtained). Beam switching (cf. § 5·3) between four aerials was employed to allow determination of both elevation and azimuth. The design and positioning of aircraft aerials of this type is an art rather than a science and we must refer the reader to the original references* for further details. The display was of the back-to-back type (cf. § 5·3) with one tube for range and azimuth determination and a second tube for elevation. Typical characteristics for this type of equipment are $G_s = 3$, $G_r = 3$, $P_s = 10^4$ W., $f_p = 1500$ c./sec., $\tau = 2\,\mu$sec., $P_m = 10^{-13}$ W., giving $r_m = 10$ km.

(2) A later form of A.I. operates on a wave-length of 10 cm. This uses a 29 in. paraboloidal aerial mounted in the nose of the aircraft. Helical scanning (cf. § 7·2) is employed, the beam being scanned from -15 to $+50°$ in elevation and through 180° in azimuth (limited by aircraft structure). The displays are of the range-azimuth and range-elevation type (cf. § 5·2·3) with arrangements for correlating the two sets of observations. Typical characteristics are $P_s = 5 \times 10^4$ W., $G_s = G_r = 400$, $\tau = 0{\cdot}5\,\mu$sec., $f_p = 1600$ c./sec., azimuthal scan speed = 360 r./min., total scan time (for one complete scan) = 2 sec., $P_m = 1{\cdot}2 \times 10^{-12}$ W., giving $r_m = 25$ km. This calculation makes no allowance for feeder losses and scanning loss factor (cf. p. 100), and in practice a range of 10 miles was considered a good performance. Compare here § 7·4, and also equations (7·1) and (7·3). Taking $\omega = 0{\cdot}05$ steradian, $\Omega = \pi$ steradian, we see that (7·1) gives $k \leqslant 50$, while (7·3) gives $k \leqslant 15$. It would therefore appear that f_p could have been increased with advantage, giving a higher k factor; the limitation due to ground returns (cf. p. 107), however, might make this undesirable.

(2*a*) An alternative form of centimetre A.I. is that using spiral conical scan. Details of this equipment are given on pp. 101–3, and fig. 7·2.

(3) H_2S is another airborne centimetre radar. Its function is to provide a map in an aircraft of the terrain beneath it (cf. § 5·4). The aerial consists of a wave guide radiator (cf. Plate III B) and a reflector

* Russell, B., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 567 (1946).

shaped to give a cosecant diagram, rotated in azimuth at about 60 r./min. The display is of the P.P.I. type, but with a non-linear time-base (cf. § 5·2·2). Typical characteristics are $\lambda = 3$ cm., $G_s = G_r = 4000$, $\tau = 1\,\mu$sec., $P_s = 5 \times 10^4$ W., $P_m = 10^{-12}$ W. and ranges of the order of 50 km. are obtained on moderate size towns.

(3*a*) The design of centimetre A.S.V. ('Air to Surface Vessels') follows closely that of the H_2S equipment, although since A.S.V. is normally used at low altitudes a slightly different design of aerial is called for (α_m being different, cf. § 5·4).

9·3. Shipborne radar

Shipborne radars operating on wave-lengths of 7 and 3 m. were developed and used during the early part of the war. Beyond the fact that the aerials used were very light and of sufficiently small dimensions to be mounted at the top of a ship's mast with no other supporting structure, these early sets had no other special features. We therefore describe some of the later developments.

(1) A typical naval 50 cm. set is the Mark II pom-pom director.* This uses two Yagi aerials, one for sending and the other for receiving, and both are mounted directly on the optical sight mounting. The radar is used as a range-finder only. Typical characteristics are $P_s = 2{\cdot}5 \times 10^4$ W., $G_s = G_r = 17$, $\tau = 2\,\mu$sec., $f_p = 500$ c./sec., and $P_m = 3 \times 10^{-12}$ W., giving $(r_m)_{\text{f.s.}} = 8$ km.

(2) An important naval centimetre radar is the 10 cm. low angle equipment. This was used during the war for surface warning by the R.A.F. and the Army from coastal sites as well as on naval vessels. No elevation-finding is provided, the beam being concentrated in the horizontal direction. When used on board ship a stabilized platform or mounting for the aerial is essential. A 'cheese' type of aerial is used with beam-switching to provide accurate bearing determination. Typical characteristics are $\lambda = 10$ cm., $P_s = 5 \times 10^5$ W., $P_m = 10^{-12}$ W., $G_s = G_r = 1200$, $\tau = 1\,\mu$sec., and $f_p = 500$ c./sec., giving $(r_m)_{\text{f.s.}} = 82$ km.

* Coales, J. F., Calpine, B. A. and Watson, D. S., *J. Instn Elect. Engrs*, vol. 93, part III A, p. 349 (1946).

Chapter 10

SECONDARY RADAR

In Chapter 1 we distinguished between primary and secondary radar, and have so far dealt only with the former. In the present chapter we discuss briefly the principles and uses of secondary radar where these differ from those already considered. It will be remembered that secondary radar is characterized by the use of a responder equipment in the target to be located. When triggered by a pulse from the radar sender, this responder generates a second pulse which is the signal observed at the radar receiver. Such a responder can generate much more energy than would be returned by simple reflexion from the target. One use of secondary radar is to give a stronger signal. It has other uses, however; it may be used, as it was in the war, to identify friendly as distinct from enemy aircraft, or by coding to provide 'personal' identification of a particular aircraft. The former system, proposed by Watson-Watt,* has become known as I.F.F. ('Identification Friend or Foe').

The whole subject of secondary radar is therefore closely linked with that of identification, and its use up to date has been largely for this purpose. Coding, whether used for personal identification or as a measure of security against enemy imitation, may take the form of an intermittent operation (coding in time) such as the emission of morse letters; in addition, the returned pulse-length may be varied for coding, or the radio-frequency of reply may be altered. It should be noted that, in general, the radio frequency of 'interrogation' (i.e. of the sender at the radar station) f_i, say, and of 'reply' (f_r) may be different, although in many systems they are the same. Identification systems are generally ancillary to an ordinary radar station (the latter being used to detect enemy craft and others without a responder fitted). Secondary radar may also be used alone to give a better signal strength, this latter system being particularly useful where 'clutter' would otherwise be a problem (cf. Chapter 8). Not only is the intensity of the radar response much increased, but by making f_i and f_r different, clutter can be eliminated completely.

* *J. Instn Elect. Engrs*, vol. 93, p. 378 (1946).

10·1. The responder

The responder consists in principle of a receiver, pulsing circuits, and a sender. The receiver is sensitive to f_i, while the sender radiates at a frequency f_r. It is important that the pulsing circuits should give the minimum delay between the receipt of the interrogating pulse and the emission of the reply, since the system is generally used for range measurement like any other radar; however, a delay that is constant at all times and the same for all responders is allowable, since a correction can then be applied. The receiver sensitivity must not be too great, or the responder will trigger spuriously; this may be a serious fault if it occurs too often, as it will fill the display with random pulses. The responder must not re-trigger due to receiving its own signals; this is usually arranged by making the system insensitive for a fixed time after the receipt of any pulse. On the other hand, the paralysis time must not be too great or many wanted pulses will be suppressed. This 'saturation' of a responder is in fact a serious limitation of secondary radar systems, and must not be overlooked. It is particularly likely to occur when the responder can be triggered by a large number of interrogating stations simultaneously. For this reason interrogation pulse repetition rates are made as low as is consistent with obtaining a satisfactory display.

The most usual type of responder consists of a super-regenerative receiver and a triggered sender. This form of apparatus allows a considerable economy of valves.* A typical receiver sensitivity used is about 5×10^{-11} W. (65 μV. across 80 ohms), which is readily achieved with a simple super-regenerative receiver with automatic gain stabilization. Responder delays less than 1 μsec. are possible with a super-regenerative receiver, so the range-error is only about 0·1 mile if, as is usually the case, the delays in the rest of the equipment are negligible. The power of the reply sender in the responder does not generally need to be very large (see examples below) so that the equipment can be quite small. Limitations of the mean dissipation of the oscillator valve may be another cause of saturation of the responder, setting an upper limit to the mean pulse frequency allowable. In practice this limit is in the region of 20,000 pulses per second.

* Wood, K. A., *J. Instn Elect. Engrs*, vol. 93, part IIIA, p. 481 (1946).

In addition to the responder equipment, an aerial is needed on the target. Single aerial working (cf. § 5·2·1) is general, except when f_i and f_r are very different. The type of aerial varies with the application, but vertical polarization is usually employed rather than horizontal. Not only does this enable a simpler aerial* to be used when sensitivity in the whole 360° of azimuth is required; it also tends to remove 'gaps' in the vertical polar diagram of the ground station, since the reflexion coefficient of the ground for vertical polarization is numerically smaller than for horizontal (cf. figs. 3·5 and 3·8). The main disadvantage of vertical polarization, viz. stronger sea clutter† (and often land clutter too) is less important with secondary radar, on account of the greater returned field-strength, and is completely absent when f_i and f_r are different.

10·2. The secondary radar system

The formulae for one-way transmission given in Chapter 3 (cf. example 1 (*a*), p. 50) apply directly to this case. Writing P_s and P_m for the ground radar sender and receiver respectively, and P'_m and P'_s for the responder receiver and sender, we have a condition for the interrogating signal relating P_s with P'_m, and another for the reply signal relating P'_s with P_m. If the wave-lengths are the same for both channels and single aerial working (or similar sending and receiving aerials) are used, we find exactly similar equations for the two directions of transmission—e.g. when ground reflexions are negligible

$$P'_m \leqslant \frac{P_s G_1 G_2 \lambda^2}{(4\pi r)^2} \quad \text{and} \quad P_m \leqslant \frac{P'_s G_1 G_2 \lambda^2}{(4\pi r)^2}, \tag{10·1}$$

where G_1 and G_2 are the gains of the radar and responder aerials respectively.

It is thus possible to design a system having

$$P_s P_m = P'_s P'_m \tag{10·2}$$

in which case the range at which the interrogator ceases to trigger the responder is the same as that at which the reply is just detectable. Such a 'balanced' design is of course advantageous. Typical values of P_m and P'_m are $2·5 \times 10^{-12}$ and 5×10^{-11} W. respectively, so that

* A vertical half-wave aerial has a circular azimuthal polar diagram, whilst a horizontal half-wave aerial has a figure of eight azimuthal diagram.

† Cf. § 8·1.

to obtain 'balance' the radar sender power should be some 20 times the responder power.

In practice, of course, ground reflexions must be taken into account by introducing a factor ϕ (cf. § 3·3·2). The formula for one-way transmission then becomes

$$P'_m \leqslant \frac{P_s G_1 G_2 \lambda^2 \phi^2}{(4\pi r)^2}. \tag{10·3}$$

ϕ is, however, no longer of the simple form $2 \sin (kh \sin \alpha)$ since $\rho \neq -1$ (cf. § 3·1·2). Further, if f_i and f_r are unequal, the values of ϕ for the two paths will in general be different. In this case a truly 'balanced' system cannot be obtained, and it has to be remembered that the range will be that of the weaker link for any particular elevation, the powers used being chosen accordingly. The same applies if different aerials are used for sending and receiving.

From example 1 (*a*) on p. 50 it is seen that adequate range is easily obtained on one-way transmission. The free-space range of one million miles obtained in this example is not typical, since responders are less sensitive than the receiver postulated, and lower sender powers or array gains are usual. However, even for systems giving much lower free-space ranges one obtains a satisfactory signal from practically any target which is above the horizon* at the radar station, and owing to diffraction the system may remain operative for some distance into the geometrical shadow zone. This is because for vertical polarization ϕ seldom falls below 0·1 or 0·2 (except at the horizon). The use of an unnecessarily high interrogator power is also undesirable, since it tends to produce saturation of the responders due to triggering from other radar stations at great distances.

10·2·1. *The Aircraft Control Radar (A.C.R.) equipment*

Radar is very useful for locating aircraft in congested traffic areas, e.g. near a landing strip or airfield. For the elimination of short-range ground clutter a secondary radar system is desirable, especially as it can also give 'personal' identification when required;

* A useful formula for calculating the effective horizon distance is

$$d = 4000 (\sqrt{h} + \sqrt{H}),$$

where h and H are the mean heights above ground of the two aerials—all distances being measured in metres.

thus the pilot of an aircraft which has been given permission to land can switch his responder off momentarily to make sure that he is being correctly identified on the display. Methods have now been developed by which not only personal identity but the height of the aircraft and other information can be transmitted along the reply channel and automatically correlated with the radar data.

The design constants of this system are:

Interrogation channel: $\lambda = 10$ cm., $P_s = 2 \times 10^5$ W., $G_1 = 2000$, $G_2 = 2$, $P'_m = 10^{-6}$ W. (a somewhat variable figure), $f_p = 1000$ c./sec., $\tau = 1\,\mu$sec., giving $(r_m)_{\text{f.s.}} = 225$ km.

Reply channel: $\lambda = 1{\cdot}5$ m., $P_s = 5$ W., $P_m = 2 \times 10^{-12}$ W., $G_1 = 1{\cdot}6$, $G_2 = 25$, giving $(r_m)_{\text{f.s.}} = 1200$ km.

The main directivity of this system is in the interrogating channel, a beamed sending array ($G = 2000$) being used. This is designed to give a cosecant law in elevation as well as very small side lobes in azimuth. The latter requirement is necessary because the directivity of the reply channel is relatively poor, so that a satisfactory response on the P.P.I. will only be obtained if the triggering of the responder is confined to the main interrogating beam. The cosecant law of the interrogator aerial ensures that the signal strength is substantially independent of range for an aircraft at constant height. This is not so on the reply channel, but this has a greater free-space range to permit satisfactory operation when ground reflexions interfere unfavourably.

10·3. The problem of identification

The earliest identification systems were I.F.F. systems intended to distinguish friendly from hostile aircraft. The ground station was the normal C.H., so that f_t and f_r were both equal to f_0, the carrier frequency of the primary radar. The C.H. stations were spread over a considerable frequency band to avoid problems of mutual interference. To enable one responder to cover the whole band, its operating frequency was continuously swept by means of a rotating condenser; on any one station the I.F.F. appeared as an intermittent signal as the responder swept through the appropriate frequency. This frequency-sweep method has been used since in many applications where one responder is required to give an intermittent signal over a wide band of frequencies. Such an inter-

mittent signal is not, however, very suitable for working with continuously rotating beamed stations, since the beam may scan across the target* in the interval between two responses. A weak response would then be missed altogether, while a stronger signal might give a return on side-lobes of the receiving aerial.† On a P.P.I. with a rotational speed of 6 r.p.m., for example, strong signals recurring once every 2·5 sec. would appear at 90° intervals, giving no real indication of the azimuth of the target, although allowing a range measurement.

For these reasons, I.F.F. systems are not generally used with rotating beam stations. An exception is with the 200 Mc./sec. G.C.I., where a rapidly chopped signal was used to give personal identification of the fighter being controlled in a particular interception. As a general policy, a completely separate identification channel was used with rotating-beam stations (cf. § 10·3·1 below).

10·3·1. *Correlation with information from the primary radar*

When a separate identification channel is used, it is necessary to arrange the displays so that the I.F.F. response and the radar response of the same aircraft can be immediately correlated. In simple cases, e.g. airborne installations, a correlation of range only may suffice (generally these radars only observe one target at once). A simple vertical rod aerial may be used for the identification channel in such a case, and the display may be of the range-amplitude type showing radar and I.F.F. signals either together or alternatively at the turn of a switch.

For ground stations larger aerials and additional displays are often allowable; it is then much easier to meet the requirements of adequate range, good azimuthal and range discrimination and good correlation with the radar. A directional aerial with beam-switching (cf. § 5·3) may be used for reception of identification signals, with a separate range-amplitude display; correlation with the P.P.I. signals may then be done in a similar manner to that used for height-finding (cf. p. 94). A double trace on which both radar and I.F.F. signals can be displayed (fig. 10·1) is of advantage for such a system,

* The target of course uses an omni-directional aerial.

† It should be noted that a side lobe 10 db. below the main beam will give a 20 db. loss on two-way transmission (i.e. primary radar echoes) but only a 10 db. loss for secondary radar signals, providing the outgoing signal is strong enough to trigger the responder at all on the side lobes.

giving range correlation at sight. Attempts to display the I.F.F. signals directly on the P.P.I. have been made, but have not been very successful.

Radar stations, using centimetric wave-lengths, often produce very narrow beams, whereas I.F.F., using wave-lengths above 1 m., can only produce comparable beams if impossibly large aerials are used. This is one of the fundamental disadvantages of I.F.F. as

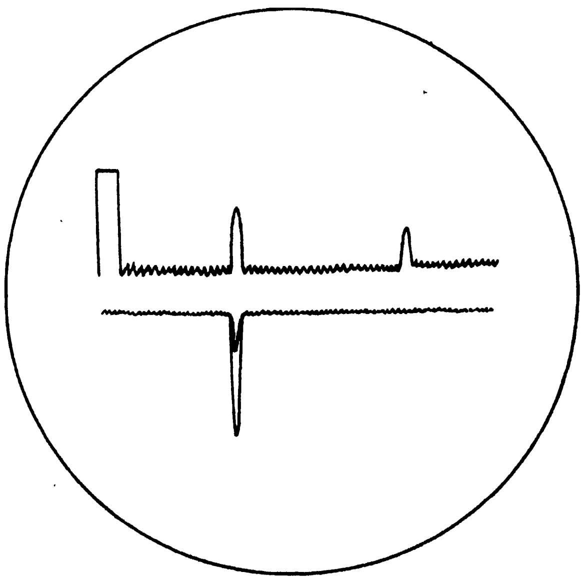


Fig. 10·1. Combined display for I.F.F. and radar. The upper trace shows radar echoes and the lower trace I.F.F. signals. Note the double response due to beam switching.

used in the past. It would seem that future identification systems should operate on a much higher frequency band, permitting directive aerials to be used even for airborne installations. For ground stations, identification signals might then be put on a P.P.I. display, especially if the responder could operate continuously on a fixed frequency instead of having to 'sweep' a band.

10·3·2. *An example of an I.F.F. system*

The system of I.F.F. developed for working on the frequency band 157–187 Mc./sec. may be taken as an example. A band of this width was necessary so that the various interrogating systems could

be staggered in frequency to avoid 'saturation' of the system by over-interrogation. The responder swept across this band once in $2\frac{1}{2}$ sec. The airborne aerial had a gain of about unity; the interrogating aerial consisted of two vertical half-wave aerials in a vertical line fed in phase (gain near the horizontal plane about 3) and the receiving aerial (using beam switching) had a gain of about 2 in the signal-equality direction. This aerial was relatively non-directive so that signals in other directions would not be missed completely.

Let us consider a case where $P_s = 10^3$ W., $P_m = 10^{-12}$ W., $P'_s = 5$ W., $P'_m = 1{\cdot}25 \times 10^{-10}$ W. (or $100\,\mu$V. across 80 ohms) and $\lambda = 1{\cdot}7$ m. For these figures we obtain from equation (10·1) a free-space range of 670 km. for interrogating and 425 km. for reply. We thus see that this system will give a certain degree of over-triggering, although as the sending and receiving aerials are different exact 'balance' in the design is not possible. When interference effects are allowed for, these results mean that a range of 85 km. is still obtainable when ϕ is as low as 0·2 (a figure usually only obtained with vertical polarization near zero elevation, cf. fig. 3·8).

Appendix 1

GLOSSARY OF SYMBOLS AND STANDARD FORMULAE

A1·1. Symbols and units

Note that all alternating quantities (e.g. current, potential, etc.) are given as r.m.s. values. Energy fluxes, etc., are all given as average values over any integral number of complete cycles.

I, i = current in amperes. V = potential in volts. v = velocity in m./sec. R = resistance in ohms. r = radial distance in metres. r_m = maximum value of r. H, h = vertical height in metres. α = elevation angle in radians. E = field strength in volts per metre. F = energy flux in field in watts/sq.m. P, P_s, P_r, P_m, etc. = power in watts. P_s = sender power. P_r = power received. P_m = limiting noise sensitivity. G = gain with respect to isotropic radiator. G' = gain with respect to half-wave aerial. Note that $G = 1{\cdot}64G'$. A, A_a, A_s, see Appendix 2. λ = wave-length in metres. f = frequency in c./sec. f_0 = carrier or mid-band frequency. f_a, f_b = upper and lower limits of frequency used. B = band width $= f_a - f_b$, or $= \frac{1}{G_0}\int_0^\infty G\,df$ (eqn. (2·1)). f_p = pulse recurrence frequency. f_i = interrogation frequency. f_r = reply frequency. τ = pulse length in sec. N = noise factor (expressed either as a pure number or as $10 \log_{10} N$ db.). Π = perception factor (cf. § 2·5). ϕ = elevation field strength factor (cf. § 3·3·2). I_p = performance index. $k = 2\pi/\lambda$, also in § 2·6 for Boltzmann's constant $= 1{\cdot}37 \times 10^{-23}$ joule/° K., and in § 7·1 for number of pulses required.

A1·2. Standard formulae

(1) $\lambda f = c$. $f = c/\lambda$.

(2) Echo delay $= 2r/c$, i.e. approximately 10·7 μsec. per mile.

(3) Maximum pulse recurrence frequency $\sim 0{\cdot}4c/r_m$ (cf. § 7·1).

(4) Pass band of receiver needed for reception of pulses is given by $B = n/\tau$, where n may be between 5 and 10 depending on how

necessary it is to preserve the steep-rising characteristic (cf. § 2·3). The range accuracy is $= \frac{1}{2}c\tau/n = c/2B$.

(5) Energy flux is given by $F = E^2/120\pi = E^2/377$ (eqn. (3·2)). Also $F = PG/4\pi r^2$ (eqn. (3·4)). Whence we have, using $G = 1{\cdot}64G'$, $E = 7\sqrt{(PG')}/r$, the usual field-strength formula. It may be noted that in the case of a half-wave aerial of radiation resistance 73 ohms carrying current I amp., $P = 73I^2$ and $G' = 1$ (defn.), whence we have $E = 60I/r$.

(6) For any aerial correctly matched into a receiver, the effective absorbing area is (cf. § A2·2)

$$A_a = \lambda^2 G/4\pi \doteqdot \lambda^2 G'/8 \quad \text{(eqn. (3·1))}.$$

(7) The power available from an aerial of gain G placed in a radiation field of strength E V./m. is

$$P_a = \frac{\lambda^2 G}{4\pi} \times \frac{E^2}{120\pi} = \frac{G\lambda^2 E^2}{480\pi^2} \quad \text{(eqn. (3·3))}.$$

(8) The formula for one-way transmission is

$$P_r = \frac{P_s G_s G_r \lambda^2}{(4\pi r)^2}.$$

(9) The formula for two-way transmission (i.e. radar) is

$$P_r = \frac{P_s G_s G_r \lambda^2 A}{64\pi^3 r^4} \quad \text{(eqn. (3·11))},$$

and therefore $$r_{\text{max.}} = \sqrt[4]{\frac{P_s G_s G_r \lambda^2 A}{64\pi^3 P_m}} \quad \text{(eqn. (3·12))}.$$

(10) The formula giving the field-strength polar diagram of a uniformly fed array is $f(\alpha) = 2\sin(\frac{1}{2}ka\sin\alpha)/k\sin\alpha$, where a is the aperture of the array (cf. § 3·1·1).

(11) The interference factor (cf. § 3·3·2) in the case of a symmetrical array used on a flat site is of the form $2\sin(hk\sin\alpha)$:

$$\phi(\alpha) = 2f(\alpha)\sin(hk\sin\alpha)$$

(12) Taking account of ground reflexions the two-way transmission formula becomes

$$(r_m)_{\text{int.}} = (r_m)_{\text{f.s.}} \times \phi(\alpha) \quad \text{(cf. eqn. (3·13))}.$$

(13) To compute α, making allowance for the curvature of the earth, we use

$$5280 \sin\alpha = h/r - \tfrac{1}{2}r.$$

Note that in this special case h is in feet and r is in miles (in all other formulae used in this book distances are measured in metres).

(14) For formulae for complex reflexion coefficient of ground see equations (3·5) and (3·6) of § 3·1·2.

(15) For formulae for range of detection of surface vessels see § 3·2·2.

(16) The limiting noise sensitivity of a receiver is given by

$$P_m = \Pi N k T_0 B, \quad \text{where} \quad T_0 = 290^\circ\,\mathrm{K}.$$

If T' is the effective aerial temperature, then effective noise factor is

$$N - 1 + T'/T_0 \quad (\text{cf. §§ 2·4·1 and 3·3·1}).$$

(17) The performance index (I_p) is given by

$$I_p = \frac{P_s G_s G_r \lambda^2}{64\pi^3 P_m}.$$

It should be noted that this is the concept preferred by the authors. There are other concepts (cf. § 3·3·3).

(18) For a two-dimensional scanning system the maximum picture repetition rate (N per sec.) is given by

$$N \leqslant \frac{f_p \omega}{k\Omega} \quad (\text{eqn. (7·1)}),$$

where ω, Ω are the solid angles of the beam and of the scan respectively, and k is a factor depending on the type of scan used (cf. § 7·1) and the accuracy of angular measurement required.

(19) For such a system the maximum range is given by

$$r_m = \sqrt[6]{\left[\frac{P_s\, 0{\cdot}4A}{P_m \pi}\left(\frac{\lambda c}{Nk\Omega}\right)^2\right]} \quad (\text{eqn. (7·4)}).$$

Appendix 2

THE CALCULATION OF ABSORBING, SCATTERING AND ECHOING AREAS

Definitions

If incident energy flux density is F_i $(= E_i^2/120\pi)$ and if A_a, A_s, and A are the equivalent absorbing, scattering and echoing areas respectively, then

$A_a F_i$ is the power 'available' (cf. § 2·4·1) at the receiver,

$A_s F_i$ is the power scattered (reradiated), and

AF_i is the power which, fed into an isotropic radiator, would reproduce the echoed intensity back along the direction of incidence.

Note that if the reradiating system has a gain G in this direction, $A = GA_s$. Also $A = 4\pi L^2$ (eqn. (3·7)).

A 2·1. Scattering and echoing by tuned half-wave parasitic element

By parasitic we understand short-circuited, so that there is no feeder and therefore no absorption. We assume the element is parallel to the E-vector of the incident field (E_i V./m.). We also assume that the aerial is resonant, i.e. of such length, very close to $\frac{1}{2}\lambda$, that its centre-point impedance is non-reactive and equal to its radiation resistance R_r. A current will be induced by the field, and since the dipole is non-reactive this current will be in phase with the field. Let its value at the centre be I, and we assume its distribution sinusoidal, so that at distance x from the centre its value is $I\cos(2\pi x/\lambda)$. The energy drawn from the field by this current will be $\int E_i I \cos\phi \, dx$ along the aerial, ϕ being the phase-angle of the current, in our case zero. This energy must be equal to the energy reradiated by the current flowing, which is I^2R_r. Therefore

$$I^2R_r = \int_{-\frac{1}{4}\lambda}^{+\frac{1}{4}\lambda} E_i I \cos\phi \cos\frac{2\pi x}{\lambda}\, dx$$

$$= E_i I \frac{\lambda}{2\pi} \int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi} \cos\phi \cos\theta \, d\theta = \frac{\lambda E_i I}{\pi} \cos\phi.$$

Since $\cos\phi = 1$, we have $I = \lambda E_i/\pi R_r$. The energy scattered is therefore $(\lambda E_i/\pi)^2/R_r$, whence we obtain

$$A_s = \frac{(\lambda E_i/\pi)^2}{R_r} \times \frac{120\pi}{E_i^2} = \frac{120\lambda^2}{\pi R_r} = 0{\cdot}52\lambda^2,$$

if we insert the usual value of 73 ohms for R_r.

To obtain the echoing area we remember that the power gain of a $\frac{1}{2}\lambda$ aerial for radiation in its equatorial plane over an isotropic source is 1·64. The echoing area is therefore $1{\cdot}64 \times 0{\cdot}52\lambda^2 = 0{\cdot}86\lambda^2$, giving $L = 0{\cdot}26\lambda$. The same result may be obtained by using the formula $60I/r$ for the field strength at distance r from a dipole carrying current I.

If the dipole is inclined at an angle θ to the electric vector, L must be multiplied by $\cos^2\theta$. If the dipole is off-tune, its impedance becomes reactive by an amount which is dependent on the cross-section of the dipole. For example, for a dipole 1 cm. in diameter and $\lambda = 500$ cm., the reactive impedance is about 10 ohms for 1 % change in length. Thus, if the dipole is 12 % above or below the correct length for resistive impedance, L will be about halved.

A 2·2. Scattering and echoing by half-wave aerial matched to a receiver

It is well known that the maximum power transfer into a load occurs when the load is equal to the impedance of the source. To obtain maximum signal into a receiver we therefore connect it to a dipole so that its impedance R_a at the centre-point feed is equal to R_r, which is assumed resistive (i.e. the dipole is resonant of length). In this case the energy absorbed from the incident field is given by the energy equation

$$I^2(R_r + R_a) = \frac{\lambda E_i I}{\pi}\cos\phi, \quad \text{so that} \quad I = \frac{\lambda E_i}{\pi(R_r + R_a)} \quad \text{if} \quad \cos\phi = 1$$

and if $R_a = R_r$ this becomes $I = \lambda E_i/2\pi R_r$, i.e. half the current in a parasitic dipole. The energy absorbed from the incident field is thus equal to half the previous amount, and half of this is re-radiated or scattered (I^2R_r), while the other half (I^2R_a) is absorbed and fed into the receiver.

Consequently the equivalent scattering area and the equivalent absorption area become respectively

$$A_s = \frac{120\lambda^2 R_r}{\pi(R_a+R_r)^2} \quad \text{and} \quad A_a = \frac{120\lambda^2 R_a}{\pi(R_a+R_r)^2},$$

so that if $R_a = R_r$ both become equal to $30\lambda^2/\pi R_r$ or approximately $\lambda^2/8$. Strictly A_a is defined in terms of 'available' power, and is therefore always equal to the latter expression, whereas A_s depends on the matching conditions applying in any given case. The equivalent echoing area is 1·64 times A_s, or $0{\cdot}21\lambda^2$, and $L = 0{\cdot}13\lambda$, approximately.

A 2·3. Absorbing area of any aerial

We have seen that for a half-wave aerial the equivalent absorbing area is approximately $\lambda^2/8$. It follows directly from the reciprocity theorem that the absorbing area of any aerial is proportional to its gain relative to the required propagation direction. Since the gain $G = 1{\cdot}64$ for the equatorial plane of a half-wave aerial, for any other aerial we have $A_a = G\lambda^2/(8\times 1{\cdot}64)$, or $G\lambda^2/4\pi$. These results are obtained approximately, but the last expression can be shown to be exact, as follows.

Consider a rectangular radiating surface (cf. para. 2 of §3·1·1) uniformly fed of area A_r, the field at the radiating surface being E_0; the total energy flux out across this surface will then be

$$A_r E_0^2/120\pi = P_s.$$

By Kirchhoff's formula the field at distance r on the normal at the centre of the area will be

$$E_r = \frac{1}{\lambda}\iint_A \frac{E_0}{r} e^{(2\pi r j)/\lambda}\, dS,$$

which for large r gives $\frac{E_0}{\lambda r}\iint_A dS$ times a phase factor, so that numerically $E_r = E_0 A_r/\lambda r$. The flux of energy at this point is therefore $\frac{E_r^2}{120\pi} = \frac{A_r^2 E_0^2}{120\pi\lambda^2 r^2} = \frac{P_s G}{4\pi r^2}$ by definition, whence, inserting the value given above for P_s, we obtain $G = 4\pi A_r/\lambda^2$. Now in the above the area A_r was supposed to radiate all its energy into space towards the observer—such an aerial will be a perfect absorber for energy

incident from the same direction. We may therefore write $A_a = A_r$, and obtain the general result

$$G = \frac{4\pi A_a}{\lambda^2} \quad \text{or} \quad A_a = G\lambda^2/4\pi.$$

By the principle of reciprocity, this result must apply to all aerials having a gain G.

A 2·4. Calculation of echoing areas of perfectly conducting objects

A 2·4·1. *Sphere*

When the radius of the sphere is small compared with the wavelength, $L = 6\pi^2 a^3/\lambda^2$ giving $A = 4{\cdot}41 a^6/\lambda^4$, so that A is inversely

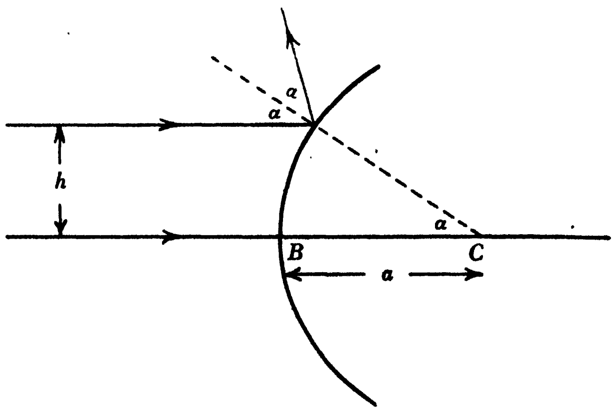


Fig. A2·1. Specular reflexion from curved surface.

proportional to λ^4. This is the well-known Rayleigh law of scattering for small targets. On the other hand, when the value of a is large compared with λ, we have $L = a/2$ (or $A = \pi a^2$). This latter result may be deduced from the principles of geometrical optics as follows.

Consider a ray (fig. A 2·1) incident on the sphere at a distance h from the ray which passes through the centre of the sphere. This ray will be deflected through an angle $\pi - 2\alpha$, where $\sin\alpha = h/a$. If unit energy crosses unit area per unit time, the amount deflected between angles $\pi - 2\alpha$, $\pi - 2(\alpha + d\alpha)$ is $2\pi h\,dh$ or $\pi a^2 \sin 2\alpha\, d\alpha$. In our notation this should be $4\pi L^2 \sin 2\alpha\, d\alpha$, whence we see that the scattering is independent of the angle of incidence and $L = \frac{1}{2}a$.

Many investigators have considered the more difficult intermediate cases when the radius of the sphere is of the same order of magnitude as the wave-length. Fig. A 2·2 gives some results calculated by Scott.* For most of the cases which arise in radar, the result $L = \frac{1}{2}a$ is sufficiently accurate. This may be seen from Scott's curve.

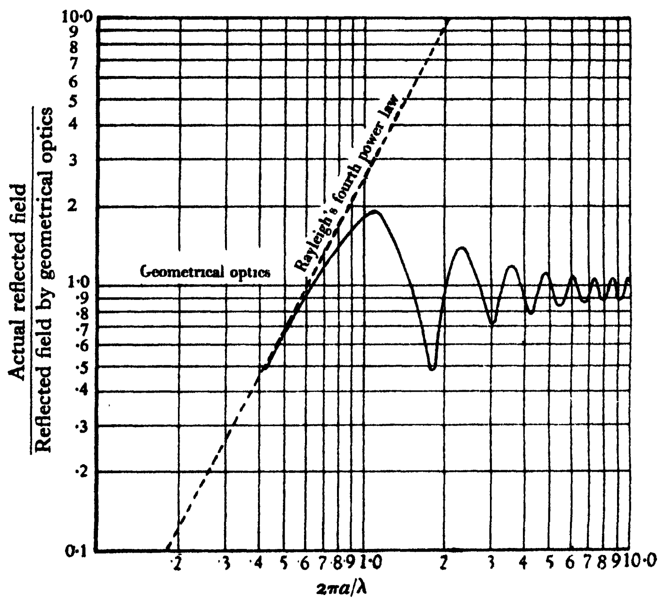


Fig. A2·2. Scattering by perfectly conducting sphere.

A 2·4·2. *Surface of any simple shape*

The back scattering from a target of dimensions large compared with the wave-length may be estimated approximately using Huygens's principle, as formulated by Kirchhoff and Fresnel. Let us suppose that radiation is emitted from a point P (fig. A 2·3) and reradiated by the surface S. Then it may be shown that, if the amplitude of the incident field at dS is E_i, the amplitude of the scattered field at P is given by

$$E_s = \frac{1}{\lambda}\int \frac{E_i e^{-j(4\pi r)/\lambda}}{r} \cos\theta \, dS,$$

* Scott, J. M. C., unpublished calculations.

where dS is any element of area and θ is the angle between the direction of the ray and the normal to dS.

Now if, as is the case in practice, r is exceedingly large compared

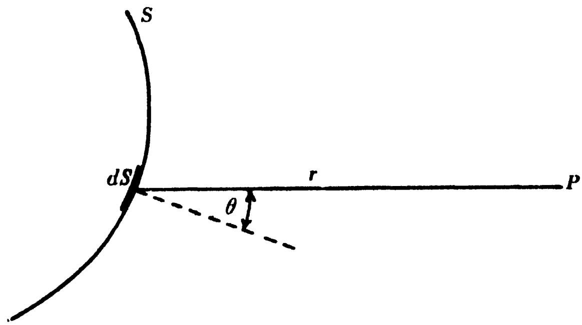


Fig. A2·3.

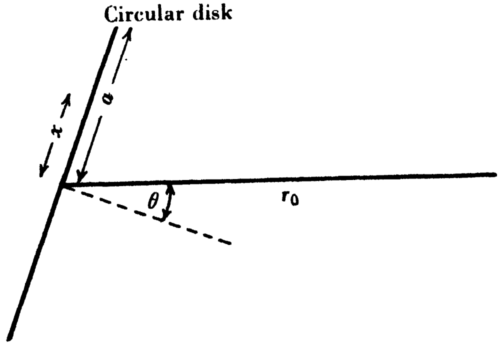


Fig. A2·4.

with the dimensions of the reflector, we can consider E_i and r constant. Therefore

$$E_s = \frac{E_i}{\lambda r} \int e^{-j(4\pi r)/\lambda} \cos\theta \, dS,$$

and thus numerically

$$L = \frac{1}{\lambda} \left| \int e^{-j(4\pi r)/\lambda} \cos\theta \, dS \right|.$$

It is usually convenient to refer the phase of the wavelet from any element dS to that from some standard (e.g. central) element

dS_0, situated at a distance, say r_0, from P. In the case of a circular disk of radius a inclined at an angle θ to the ray direction (fig. A 2·4) we have

$$L = \frac{\cos\theta}{\lambda}\int_{-a}^{+a} 2\sqrt{(a^2-x^2)}\, e^{-j(4\pi x \sin\theta)/\lambda}\, dx.$$

Hence
$$L = \frac{a\cot\theta}{2} J_1\left(\frac{4\pi a \sin\theta}{\lambda}\right).$$

For normal incidence this reduces to $L = \pi a^2/\lambda$, and gives $A = 4\pi^3 a^4/\lambda^2$.

INDEX

DATE OF ISSUE

This book must be returned within 3, 7, 14 days of its issue. A fine of ONE ANNA per day will be charged if the book is overdue.